Empirical Welfare Maximization with Constraints*

Liyang Sun[†]

August 30, 2021

Abstract

When designing eligibility criteria for welfare programs, policymakers naturally want to target the individuals who will benefit the most. This paper proposes two new econometric approaches to selecting an optimal eligibility criterion when individuals' costs to the program are unknown and need to be estimated. One is designed to achieve the highest benefit possible while satisfying a budget constraint with high probability. The other is designed to optimally trade off the benefit and the cost from violating the budget constraint. The setting I consider extends the previous literature on Empirical Welfare Maximization by allowing for uncertainty in estimating the budget needed to implement the criterion, in addition to its benefit. Consequently, my approaches improve the existing approach as they can be applied to settings with imperfect take-up or varying program needs. I illustrate my approaches empirically by deriving an optimal budget-constrained Medicaid expansion in the US.

Keywords: empirical risk minimization, heterogeneous treatment effects and costs, cost-benefit analysis, individualized treatment rules

JEL classification codes: C14, C44, C52

1 Introduction

When a welfare program induces varying benefits across individuals, and when resources are scarce, policymakers naturally want to prioritize eligibility to individuals who will benefit the most. Based on experimental data, cost-benefit analysis can inform policymakers on which subpopulations to prioritize, but these subpopulations might not align with any available eligibility criterion such as an income threshold. Kitagawa and Tetenov (2018) propose Empirical Welfare Maximization (EWM) that can directly select an eligibility criterion from a set of available criteria. For example, if available criteria take the form of income thresholds, EWM considers the problem of maximizing the expected benefits in the population

$$\max_{t \leq \bar{t}} \ E[benefit \cdot \mathbf{1}\{income \leq t\}]$$

and approximates the optimal threshold based on benefits estimated from experimental data. Recent papers have studied the statistical properties of the EWM approach, including Athey and Wager (2020), among others. Notably the average benefits under the eligibility criterion selected by EWM converges to the highest attainable benefits as the sample size grows.

In practice policymakers often face budget constraints, but only have imperfect information about whether a given eligibility criterion satisfies the budget constraint. First, there may be imperfect take-up: eligible individuals might not participate in the welfare program, resulting in zero cost to the government, e.g. Finkelstein and Notowidigdo (2019). Second, costs incurred by eligible individuals who participate in the welfare program may vary considerably, largely driven by individuals' different needs but also many other factors, e.g. Finkelstein et al. (2017). Both considerations are hard to predict ex ante, implying that the potential cost of providing eligibility to any given individual is unknown at the time of designing the eligibility criterion. Unobservability of the potential cost requires estimation based on experimental data,

contributing to uncertainty in the budget estimate of a given eligibility criterion. The existing EWM approach takes cost as known, so that in the aforementioned example the upper bound $\bar{t}$ for the income threshold that satisfies the budget constraint is known. Therefore, the existing approach is not able to address issues such as a potential cost overrun due to an underestimate of the budget.

In this paper I provide two new econometric approaches to selecting an optimal eligibility criterion when the budget needed to implement the eligibility criterion involves an unknown cost and need to be estimated. Both approaches take the form of a statistical rule, which selects an eligibility criterion based on the experimental data. The two statistical rules I propose augment the EWM literature in settings of unknown cost, each satisfying one of the two desirable properties I describe below. The first statistical rule I propose, the *mistake-controlling rule*, achieves the highest benefit possible while satisfying a budget constraint with high probability. Therefore, with high probability, the mistake-controlling rule selects only feasible eligibility criteria that satisfy the budget constraint. The second statistical rule I propose, the *trade-off rule*, can select infeasible eligibility criteria, but only if borrowing money and exceeding the budget constraint are not too costly to justify the marginal gain in benefit from violating the budget constraint. The choice between the two rules should align with policymakers' attitude toward the budget constraint. If policymakers are financially conservative, then the eligibility criterion selected by the mistake-controlling rule is more suitable. If policymakers want to reach as many individuals as they can, then the eligibility criterion selected by the trade-off rule is more suitable.

In this novel setting where the budget needed to implement an eligibility criterion involves an unknown cost, I evaluate statistical rules based on two properties: *asymptotic feasibility* and *asymptotic optimality*. A statistical rule is asymptotically feasible if given a large enough sample size, the statistical rule is very likely to select feasible eligibility criteria that satisfy the budget constraint. A statistical rule is asymptotically optimal if given a large enough sample size, the statistical rule is very

likely to select eligibility criteria that achieve weakly higher benefits than any feasible criterion. A similar notion of asymptotic optimality has been studied by the existing EWM literature. The existing EWM statistical rule is proven to be asymptotically optimal in settings where cost is known. Departing from the existing EWM literature, I focus on settings with unknown cost. Correspondingly, I add asymptotic feasibility to asymptotic optimality as desirable properties for statistical rules. Ideally, the statistical rule should maintain good performance even when the data distribution is unknown. I therefore define *uniform asymptotic optimality* and *uniform asymptotic feasibility*, which impose *asymptotic optimality* and *asymptotic feasibility* uniformly over a class of reasonable data distributions, respectively. However, the setting with unknown cost can be too complicated for an arbitrary draw of the data to approximate, and I prove that it is impossible for any statistical rule to satisfy these two uniform properties simultaneously. Since achieving simultaneous uniform asymptotic feasibility and uniform asymptotic optimality is impossible, I instead aim to achieve one at a time.

The obvious extension of the existing EWM statistical rule to settings with unknown cost is the plug-in approach which imposes an estimated budget constraint. I prove that this extension is neither uniformly asymptotically optimal nor uniformly asymptotically feasible. Hence, the direct extension of the existing EWM statistical rule is not appealing in this setting, and in this paper I seek alternative statistical rules.

The first statistical rule I propose, the mistake-controlling rule, is uniformly asymptotically feasible, meaning that given a large enough sample size, it is likely to select feasible eligibility criteria that satisfy the budget constraint regardless of the underlying data distribution. The second statistical rule I propose, the trade-off rule, can select infeasible eligibility criteria under any data distribution, but only to trade off the cost of exceeding the budget constraint against gains in the benefits. I prove this statistical rule is uniformly asymptotically optimal. These two alternative

statistical rules I propose can also be applied to settings beyond budget-constrained welfare programs. For example, problems of maximizing the accuracy of any predictive algorithm while controlling the disparity in mistakes made across different ethnic groups have the same mathematical structure and fall under the framework of this paper.

To illustrate the statistical rules proposed in this paper, I consider an optimal budget-constrained Medicaid expansion. Medicaid is a government sponsored health insurance program intended for the low-income population in the United States. Based on the Oregon Health Insurance Experiment (OHIE) conducted in 2008, I derive Medicaid expansion criteria using the two statistical rules I propose. The current Medicaid expansion criterion in Oregon is based on a threshold for household income only. In my application, I examine whether health can be improved by allowing the income threshold to vary by number of children in the household, setting the budget constraint to the current level. The mistake-controlling rule selects an eligibility criterion that limits eligibility to be lower than the current level. This lower level occurs because Medicaid costs vary considerably across households, making it harder to verify whether households meet the budget constraint. In contrast, the trade-off rule selects an eligibility criterion that expands eligibility for many households above the current level, especially those with children. The higher level occurs because Medicaid in general improves health for these households, and the additional health benefit from violating the budget constraint exceeds the cost of doing so, assuming reasonable upper bound on the monetary value on the health benefit.

The rest of the paper proceeds as follows. Section 1.1 discusses related work in more detail. Section 2 presents theoretical results. Sections 3 and 4 propose two statistical rules for selecting eligibility criteria. Section 5 conducts a simulation study to illustrate the asymptotic properties of the statistical rules I propose. Section 6 considers an empirical example of designing a more flexible Medicaid expansion criterion for the low-income population in Oregon. Section 7 concludes.

1.1 Literature review

This paper is related to the traditional literature on cost-benefit analysis, e.g. Dhailiwal et al. (2013), and to the recent literature on EWM, e.g. Kitagawa and Tetenov (2018), Mbakop and Tabord-Meehan (2019), Rai (2019) and Athey and Wager (2020). More broadly, this paper contributes to a growing literature on statistical rules in econometrics, including Manski (2004), Dehejia (2005), Hirano and Porter (2009), Stoye (2009), Chamberlain (2011), Bhattacharya and Dupas (2012), Demirer et al. (2019) and Kasy and Sautmann (2020).

The traditional cost-benefit analysis compares the cost and benefit of a given welfare program. The effect of program eligibility is first estimated based on a randomized control trial (RCT), and then converted to a monetary benefit for calculating the cost-benefit ratio. For example, Gelber et al. (2016) and Heller et al. (2017) compare the efficiency of various crime prevention programs based on their cost-benefit ratios. However, the cost-benefit ratio is only informative for whether this welfare program should be implemented with the fixed eligibility criterion as implemented in the RCT.

The literature on statistical rules in econometrics has also developed a definition for optimality of statistical rules. Manski (2004) considers the minimax-regret, defined to be loss in expected welfare achieved by the statistical rule relative to the welfare achieved by the theoretically optimal eligibility criterion. In the absence of any constraint, under the theoretically optimal eligibility criterion, anyone with positive benefit from the welfare program would be assigned with eligibility. The minimax regret is the upper bound on the regret that results from not knowing the data distribution. Manski (2004) argues a statistical rule is preferred if its minimax-regret converges to zero with the sample size, and analyzes the minimax-regret for the class of statistical rules that only selects eligibility criterion based on subsets of the observed covariates. Stoye (2009) shows that with continuous covariates and no functional form restrictions on the set of criteria, minimax regret does not converge

to zero with the sample size because the theoretically optimal criterion can be too difficult to approximate by statistical rule. Kitagawa and Tetenov (2018) avoid this issue by imposing functional form restrictions. They propose the EWM rule, which starts with functional form restrictions on the class of available criteria, and then selects the criterion with the highest estimated benefit (empirical welfare) based on an RCT sample. They prove the optimality of EWM in the sense that its regret converges to zero at the minimax rate. Importantly, the regret is defined to be loss in expected welfare relative to the maximum achievable welfare in the constrained class, which avoids the negative results of Stoye (2009). Athey and Wager (2020) propose doubly-robust estimation of the average benefit, which leads to an optimal rule even with quasi-experimental data. Mbakop and Tabord-Meehan (2019) propose a Penalized Welfare Maximization assignment rule which relaxes restrictions of the criterion class.

The existing EWM literature has not addressed budget constraints with an unknown cost. Kitagawa and Tetenov (2018) consider a capacity constraint, which they enforce using random rationing. Random rationing is not ideal as it uses the limited resource less efficiently than accounting for the cost of providing the welfare program to an individual. When there is no restriction on the functional form of the eligibility criterion, Bhattacharya and Dupas (2012) demonstrate that given a capacity constraint, the optimal eligibility criterion is based on a threshold on the benefit of the welfare program to an individual. When the cost of providing the welfare program to an individual is heterogeneous, however, budget constraints can be more complicated than capacity constraints, and require estimation. Sun et al. (2021) propose a framework for estimating the optimal rule under a budget constraint when there is no functional form restriction. The theoretical contribution of this paper is to extend the literature by allowing both functional form restrictions and budget constraints with an unknown cost.

2 Theoretical results

The policymaker's goal is to solve a constrained optimization problem as explained in Section 2.1, which also introduces two important settings that share this mathematical structure: welfare program eligibility criterion under budget constraints and algorithmic predictions under fairness constraints. In this paper I focus on the problem of designing welfare program eligibility criterion under budget constraints, in particular when the cost of providing eligibility to any given individual is unknown ex ante. Section 2.2 explains how based on experimental data, statistical rules can make progress on solving the constrained optimization problem. Section 2.3 defines two desirable properties of statistical rules. Sections 2.4 and 2.5 derive some negative implications of these difficulties on finding statistical rules with good properties.

2.1 Motivation and setup

I begin by setting up a general constrained optimization problem, which depends on the following random attributes of an individual:

$$A = (\Gamma, R, X) \in \mathcal{A} \subseteq \mathbb{R}^{2+p}. \tag{1}$$

Here Γ is her benefit from the treatment, R is the cost to the policymaker of providing her with the treatment, and $X \in \mathcal{X} \subset \mathbb{R}^p$ denotes her p-dimensional characteristics. The individual belongs to a population that can be characterized by the joint distribution P on her random attributes A. The unknown distribution $P \in \mathcal{P}$ is from a class of distributions $\mathcal{P}$.

A policy $g(X) \in \{0, 1\}$ determines the treatment status for an individual with observed characteristics X, where 1 is treatment and 0 is no treatment. Let $\mathcal{G}$ denote the class of policies policymakers can choose from. The optimization problem is to

find a policy with maximal benefit while subject to a constraint on its cost:[1]

$$\max_{g \in \mathcal{G}} E_P[\Gamma \cdot g(X)] \ s.t. \ E_P[R \cdot g(X)] \leq k. \tag{2}$$

If the policymaker does not have a fixed budget but still want to account for cost, the scalar Γ can be the difference in benefit and cost. Since a fixed budget is common, I impose a harsh constraint.

The benefit-cost attributes (Γ, R) of any given individual may or may not be observed. When the benefit-cost attributes are observed, policymakers observe i.i.d. $A_i = (\Gamma_i, R_i, X_i)$ in a random sample of the target population.[2] When the benefit-cost attributes are unobserved, policymakers observe their estimates (Γ_i^*, R_i^*) in a random sample along with the characteristics X_i. The focus of this paper is the setting where policymakers only observe (Γ_i^*, R_i^*, X_i) in a random sample from an experiment or quasi-experiment. Observing only the estimates for benefit-cost attributes introduces additional difficulties as explained in Section 2.2.

In the example of implementing welfare programs, policies take the form of eligibility criteria. In the example of predicting future behavior, policies take the form of risk assessment. I restrict attention to non-randomized policies as in the leading example of welfare programs, deterministic policies such as income thresholds are more relevant. While the optimization problem policymakers face in practice can take on different forms, constrained optimization problems of this mathematical structural are ubiquitous in social science. I next present two examples.

Example 2.1. Welfare program eligibility criterion under budget constraint

Suppose the government wants to implement some welfare program. The treatment in this example is eligibility for such welfare program. Due to a limited budget, the government cannot make eligibility universal and can only provide eligibility to a

---

[1]Following Kitagawa and Tetenov (2018), I implicitly assume the maximizer exists in $\mathcal{G}$ with the notation in (2).

[2]I use the subscript i to index individuals in this random sample.

subpopulation. To use the budget efficiently, policymakers consider the constrained optimization problem (2). In this example, the policy $g(X)$ assigns an individual to eligibility based on her observed characteristics X, and is usually referred to as an eligibility criterion. I denote Γ to be the benefit experienced by an individual after receiving eligibility for the welfare program. Specifically, let (Y_1, Y_0) denote the potential outcomes that would have been observed if an individual were assigned with and without eligibility, respectively. The benefit from eligibility criterion is therefore defined as $\Gamma := Y_1 - Y_0$. Note that maximizing benefit is equivalent to maximizing the outcomes (welfare) under the utilitarian social welfare function: $E_P[Y_1 \cdot g(X) + Y_0(1 - g(X))]$. I denote R to be the potential cost from providing an individual with eligibility for the welfare program. Both Γ and R are unobserved at the time of assignment and will need to be estimated.

Policymakers might be interested in multiple outcomes for an in-kind transfer program. Hendren and Sprung-Keyser (2020) capture benefits by the willingness to pay (WTP). Assuming eligible individuals make optimal choices across multiple outcomes, the envelope theorem allows policymakers to focus on benefit in terms of one particular outcome. $\triangle$

Example 2.2. Algorithmic Predictions under Fairness Constraints

There are many machine learning algorithms that predict individuals' future behavior such as loan repayment and recidivism. The treatment in this example is the prediction, and the policy $g(X)$ assigns an individual to be high risk for such future behavior based on her observed characteristic X. These algorithms can, however, exhibit undesirable biases. For example, the COMPAS algorithm, provided by a proprietary software company, predicts whether a defendant is high risk of recommitting crimes. Several U.S. courts use predictions from this algorithm to help assess the likelihood of recidivism for a given defendant. A ProPublica investigation by Angwin et al. (2016) showed this algorithm makes false positive mistakes unfairly among black defendants. Specifically, the actual rate of recidivism is often lower than predicted for

black defendants, whereas higher than predicted for white defendants. This disparity is termed false positive disparity by Kearns et al. (2018).

The COMPAS algorithm is proprietary, so policymakers do not know its underlying model to remove its discriminatory features. Given data on recidivism, however, policymakers may instead consider the constrained optimization problem (2), to construct a rule that follows the prediction of the COMPAS algorithm as much as possible while constraining it to make false positive mistakes with comparable frequencies for black defendants and white defendants. The benefit Γ in this optimization problem is the negative of the loss when a given policy $g(X)$ deviates from the COMPAS prediction. The cost R in this optimization problem is the *excess actual* risk for blacks relative to whites but with the same characteristics X.[3] To constrain the false positive disparity, that the discriminated group has less false negative mistake than the other group, policymakers set $k = 0$. While the loss Γ is observed because policymakers observe the prediction by the COMPAS algorithm, the false positive disparity R is unobserved as it depends on future recidivism outcomes and needs to be approximated. $\triangle$

I further illustrate the problem of selecting eligibility criterion for welfare programs with the example of Medicaid expansion.

---

[3]Formally, policymakers solve a constrained minimization problem. For any symmetric loss function $l(\cdot, \cdot)$, policymakers want to minimize the expected loss for a given policy $g(X)$ to deviate from the COMPAS prediction Y i.e. when $Y \neq g(X)$:

$$
\begin{aligned}
E_{XY}[l(Y, g(X))] =& E_X E_{Y|X}[l(Y, g(X))] \\
=& E_X[\eta(X)l(1,0)(1 - g(X))) + (1 - \eta(X))l(0,1)g(X))] \\
\propto& E_X[(1 - 2\eta(X))g(X))]
\end{aligned}
$$

where $\eta(X)$ is the (known) average prediction by the COMPAS algorithm for individuals with characteristics X. This implies $\Gamma = 1 - 2\eta(X)$. The false positive mistake happens when $g(X)$ predicts an individual to be of high risk, but this individual does not recommit crime in the future $Y^* = 0$. The excess actual risk of blacks ($S = 1$) relative to whites ($S = 0$), holding the observed characteristics constant, is defined to be

$$
R = Pr_P\{Y^* = 0 \mid S = 1, X\} - Pr_P\{Y^* = 0 \mid S = 0, X\}.
$$

Example - Medicaid Expansion Medicaid is a government-sponsored health insurance program intended for the low-income population in the United States. Up till 2011, many states provided Medicaid eligibility to able-bodied adults with income up to 100% of the federal poverty level. The 2011 Affordable Care Act (ACA) provided resources for states to expand Medicaid eligibility for all adults with income up to 138% of the federal poverty level starting in 2014.

Suppose policymakers want to maximize the health benefit of Medicaid by adopting a more flexible expansion criterion. Specifically, the more flexibility criterion relaxes the uniform income threshold of 138% and allows the income thresholds to vary with number of children in the household. It may be infeasible to implement a new expansion criterion if it costs more than the current one. Therefore I impose the constraint that the more flexible criterion must cost no more than the current one.

Correspondingly, the constrained optimization problem (2) sets Γ to be the health benefit from Medicaid, R to be the *excess* per-enrollee cost of Medicaid relative to the current expansion criterion, and the appropriate threshold to be $k = 0$. The characteristics X include both income and number of children in the household. $\triangle$

Applying the Law of Iterated Expectation, the constrained optimization problem (2) can be written as

$$\max_{g \in \mathcal{G}} E_P[E_P[\Gamma \mid X] \cdot g(X)] \ s.t. \ E_P[E_P[R \mid X] \cdot g(X)] \leq k. \tag{3}$$

When the eligibility criterion can be based on any characteristics whatsoever, the class of available criteria is unrestricted i.e. $\mathcal{G} = 2^{\mathcal{X}}$. In this unrestricted class, when the cost is non-negative, the above expression makes clear that the optimal eligibility criterion is based on thresholding by the benefit-cost ratio $E_P[\Gamma \mid X]/E_P[R \mid X]$ where the numerator and the denominator are respectively the average effects conditional on the observed characteristics (CATE) and the conditional average resource required. Appendix B provides a formal statement.

Given a random sample, to approximate the optimal eligibility criterion g_P^*, one

can estimate the benefit-cost ratio based on the estimated CATE and the estimated conditional average resource required. The resulting statistical rule selects eligibility criteria that are thresholds based on the estimated benefit-cost ratio. The challenge is that the selected eligibility criterion can be hard to implement when the estimated benefit-cost ratio is a complicated function of X. Restrictions on the criterion class $\mathcal{G}$ address this issue. A common restriction is to consider thresholds based directly on X, e.g. assigning eligibility when an individual's income is below a certain value.

Restrictions on the criterion class $\mathcal{G}$ mean that there might not be closed-form solutions to the population problem (2). In particular, the constrained optimal eligibility criterion g_P^* might not be an explicit function of the CATE and the conditional average resource required. Therefore it might be difficult to directly approximate g_P^* based on the estimated CATE and the estimated conditional average resource required. However, this is not an obstacle to deriving the statistical rules I propose. As I demonstrate later, the derivation does not require the knowledge of the functional form of the constrained optimal eligibility criterion g_P^*.

Example - Medicaid expansion (cont'd)   The criterion class in this example includes income thresholds that can vary with the number of children in the household:

$$
\mathcal{G} = \left\{ g(x) = \begin{cases} \mathbf{1}\{income \leq \beta_1\} & , \ numchild = 1 \\ \vdots & \\ \mathbf{1}\{income \leq \beta_j\} & , \ numchild = j \end{cases} \right\} \tag{4}
$$

for characteristics $x = (income, \ numchild)$ and $\beta_j \geq 0$. $\triangle$

2.2 Estimates for the welfare function and the budget function

To simplify the notation, I define the *welfare* function and the *budget* function:

$$W(g; P) = E_P[\Gamma \cdot g(X)], \ B(g; P) = E_P[R \cdot g(X)]. \tag{5}$$

As explained in Example 2.1 from Section 2.2, under a utilitarian social welfare function, maximizing the benefit with respect to eligibility criterion is equivalent to maximizing the welfare, which is why I refer to $W(g; P)$ as the welfare function. The welfare function and the budget function are both deterministic functions from $\mathcal{G} \to \mathbb{R}$. The index by the distribution P highlights that welfare and budget of criterion g vary with P, and in particular, whether a criterion g satisfies the budget constraint depends on which distribution P is of interest.

When the benefit-cost attributes (Γ, R) are unobserved and the distribution P is unknown, both the welfare function and the budget function are unknown functions. Given experimental data, using estimates (Γ_i^*, R_i^*) described below, one can estimate the welfare function and the budget function using their sample-analog versions:

$$\widehat{W}_n(g) := \frac{1}{n} \sum_i \Gamma_i^* \cdot g(X_i), \ \widehat{B}_n(g) := \frac{1}{n} \sum_i R_i^* \cdot g(X_i). \tag{6}$$

Solving the constrained optimization problem (2) based on these sample-analogs means the solution is a statistical rule $\widehat{g}$, which selects an eligibility criterion based on the observed data. Therefore, the estimation errors in (Γ_i^*, R_i^*) need to be asymptotically unimportant for the performance of statistical rules.

The appropriate expressions for these estimates depend on the type of observed data. Below I state the estimates formed based an RCT that randomly assigns the eligibility. The observed data $\{A_i^*\}_{i=1}^n$ consists of i.i.d. observations $A_i^* = (Y_i, Z_i, D_i, X_i) \in \mathcal{A}^*$. The distribution of A_i^* is induced by the distribution of (Γ, R, X) as in the population, as well as the sampling design of the RCT. Here D_i is an in-

dicator for being in the eligibility arm of the RCT, Y_i is the observed outcome and Z_i is the observed cost of providing eligibility to an individual participating in the RCT. The observed cost is mechanically zero if an individual is not randomized into the eligibility arm. The estimates for (Γ, R) are

$$\Gamma_i^* = \alpha(X_i, D_i) \cdot Y_i, \ R_i^* = \frac{D_i}{p(X_i)} \cdot Z_i, \tag{7}$$

where $\alpha(X_i, D_i) = \frac{D_i}{p(X_i)} - \frac{1-D_i}{1-p(X_i)}$ and $p(X_i)$ is the propensity score, the probability of receiving eligibility conditional on the observed characteristics. Since the sampling design of an RCT is known, the propensity score is a known function of the observed characteristics. To see why these are good estimates, note that the population problem (2) has the same solution replacing (Γ, R) with (Γ_i^*, R_i^*). Specifically, using the potential outcome notation introduced in Example 2.1 from Section 2.2, one can express (Y_i, Z_i) as $Y_i = D_i \cdot Y_{1i} + (1 - D_i) \cdot Y_{0i}$ and $Z_i = D_i \cdot R_i$. These expressions make clear that (Γ_i^*, R_i^*) are conditionally unbiased estimators for $E_P[\Gamma \mid X]$ and $E_P[R \mid X]$, and the population problem is unchanged replacing (Γ, R) with (Γ_i^*, R_i^*) as

$$E_P[\Gamma_i^* \cdot g(X_i)] = E_P[E_P[\Gamma_i^* \mid X_i] \cdot g(X_i)] = E_P[E_P[\Gamma_i \mid X_i] \cdot g(X_i)] = E_P[\Gamma_i \cdot g(X_i)]$$
$$E_P[R_i^* \cdot g(X_i)] = E_P[E_P[R_i^* \mid X_i] \cdot g(X_i)] = E_P[E_P[R_i \mid X_i] \cdot g(X_i)] = E_P[R_i \cdot g(X_i)]$$

by the Law of Iterated Expectation.

2.3 Desirable properties for statistical rules

Denote by $\widehat{g}$ a statistical rule that selects an eligibility criterion after observing some experimental data of sample size n. This section provides formal definitions for two desirable properties of $\widehat{g}$.

Definition 2.1. A statistical rule $\widehat{g}$ is *pointwise asymptotically optimal* under the

data distribution P if for any $\epsilon > 0$

$$\lim_{n \to \infty} Pr_{P^n} \{W(\widehat{g}; P) - W(g_P^*; P) < -\epsilon\} = 0, \tag{8}$$

and *uniformly asymptotically optimal* over the class of distributions $\mathcal{P}$ if for any $\epsilon > 0$

$$\lim_{n \to \infty} \sup_{P \in \mathcal{P}} Pr_{P^n} \{W(\widehat{g}; P) - W(g_P^*; P) < -\epsilon\} = 0.$$

A statistical rule is *pointwise asymptotically feasible* under the data distribution P if

$$\lim_{n \to \infty} Pr_{P^n} \{B(\widehat{g}; P) > k\} = 0,$$

and *uniformly asymptotically feasible* over the class of distributions $\mathcal{P}$ if

$$\lim_{n \to \infty} \sup_{P \in \mathcal{P}} Pr_{P^n} \{B(\widehat{g}; P) > k\} = 0. \tag{9}$$

∎

The above two properties build on the existing EWM literature. For the first property, the current EWM literature considers a statistical rule "asymptotically optimal" if it selects eligibility criteria that attain the same value as g_P^* in expectation over repeated sample draws as $n \to \infty$. Instead, I strengthen the convergence in mean to convergence in probability, where the probability of $\widehat{g}$ selecting eligibility criteria that achieve strictly lower welfare than g_P^* approaches zero as $n \to \infty$.[4] In the setting of the existing EWM literature, the constrained optimal criterion g_P^* is also the unconstrained optimal criterion $\mathcal{G}$. Therefore it is impossible for any statistical rule $\widehat{g}$ to select a criterion that achieves higher value than g_P^*. In my setting, however, the constrained optimal criterion g_P^* is not necessarily the unconstrained op-

---

[4]A more precise definition replaces the probability in (8) with $Pr_{P^n} \{\omega : W(\widehat{g}(\omega); P) - W(g_P^*; P) < -\epsilon\}$, the probability over repeated samples drawn from P^n that $\widehat{g}$ selects eligibility criteria that achieves welfare strictly less than maximum feasible level in the population distributed according to P.

timal criterion. Thus, I allow the statistical rule $\widehat{g}$ to select a criterion that achieves higher welfare than g_P^* for all data distributions, albeit only at the cost of violating the budget constraint.

The second property is new to the EWM literature. It imposes that given a large enough sample size, the statistical rule $\widehat{g}$ is unlikely to select infeasible eligibility criteria that violate the budget constraint, so that it is "asymptotically feasible".[5] Asymptotic feasibility of statistical rules is specific to the current setting where the budget constraint involves unknown cost.

While both are desirable properties, the next section shows a negative result that it is impossible for a statistical rule to satisfy both properties when the data distribution is unknown and belongs to a sufficiently rich class of distributions $\mathcal{P}$.

2.4 Impossibility result

Theorem 2.1 presents an impossibility result that no statistical rule can be both uniformly asymptotically optimal and uniformly asymptotically feasible in a sufficiently rich class of distributions $\mathcal{P}$. Assumption 2.1 explains the notion of richness: the sets of feasible eligibility criteria differ only marginally at nearby pairs of distributions. Assumptions 2.2 and 2.3 characterize distributions where such richness can be problematic for statistical rules to simultaneously achieve uniform asymptotic optimality and uniform asymptotic feasibility.

Assumption 2.1. Contiguity. *There exists a distribution $P_0 \in \mathcal{P}$ under which a non-empty set of eligibility criteria satisfies the constraint exactly $\mathcal{G}_0 = \{g : B(g; P_0) = k\}$. Furthermore, the class of distributions $\mathcal{P}$ includes a sequence of data distributions $\{P_{h_n}\}$ contiguous to P_0, under which for all $g \in \mathcal{G}_0$, there exists some $C > 0$ such that*

$$\sqrt{n} \cdot (B(g; P_{h_n}) - k) > C.$$

---

[5]A more precise definition replaces the probability in (9) with $Pr_{P^n}\{\omega : B(\widehat{g}(\omega); P) > k\}$, the probability over repeated samples drawn from P^n that $\widehat{g}$ selects eligibility criteria that violate the budget constraint in the population distributed according to P.

Assumption 2.2. Binding constraint. *Under the data distribution P_0, the constraint is satisfied exactly at the constrained optimum i.e. $B(g_P^*; P_0) = k$.*

Assumption 2.3. Separation. *Under the data distribution P_0, $\exists \epsilon > 0$ such that for any feasible criterion g, whenever*

$$\left| W(g; P_0) - W(g_{P_0}^*; P_0) \right| < \epsilon,$$

we have

$$\left| B(g; P_0) - B(g_{P_0}^*; P_0) \right| = 0.$$

Theorem 2.1. *Suppose Assumption 2.1 holds for the class of data distributions $\mathcal{P}$. For $P_0 \in \mathcal{P}$ considered in Assumption 2.1, suppose it also satisfies Assumptions 2.2 and 2.3. Then no statistical rule can be both uniformly asymptotically optimal and uniformly asymptotically feasible. In particular, if a statistical rule $\widehat{g}$ is pointwise asymptotically optimal and pointwise asymptotically feasible under P_0, then it is not uniformly asymptotically feasible.*

Remark 2.1. Figure 2.1 provides some intuition for Theorem 2.1. The pictured distribution P_0 satisfies both Assumptions 2.2 and 2.3: the constrained optimal criterion $g_{P_0}^*$ satisfies the budget constraint exactly, i.e. $B(g_{P_0}^*; P_0) = k$, and is separated from the rest of feasible eligibility criteria. Note that if a statistical rule $\widehat{g}$ is pointwise asymptotically optimal and pointwise asymptotically feasible under P_0, then it has to select eligibility criteria close to $g_{P_0}^*$ with high probability over repeated sample draws distributed according to P_0^n as $n \to \infty$.

Under Assumption 2.1, the class of distributions $\mathcal{P}$ is sufficiently rich so that along a sequence of data distributions $\{P_{h_n}\}$ that is contiguous to P_0 as $n \to \infty$, the budget functions $B(g; P_{h_n})$ converge to $B(g; P_0)$ while $B(g_{P_0}^*; P_{h_n}) > k$, i.e. $g_{P_0}^*$ is not feasible under P_{h_n}. Figure 2.1 showcases P_1 as one distribution from this sequence. The contiguity between $\{P_{h_n}\}$ and P_0 implies that the statistical rule $\widehat{g}$ must select criteria close to $g_{P_0}^*$ with high probability under $P_{h_n}^n$ as well. However, the criterion $g_{P_0}^*$

is infeasible under P_{h_n}, and therefore the statistical rule $\widehat{g}$ cannot be asymptotically feasible under P_{h_n}. $\triangle$

In Appendix B.2, I give more primitive assumptions under which Assumption 2.1 is guaranteed to hold. These primitive assumptions are relatively weak. Furthermore, it is not implausible for real-world distributions to satisfy both Assumptions 2.2 and 2.3. This suggests the impossibility results are relevant in real-world settings. For example, consider a one-dimensional criterion class $\mathcal{G}$, e.g. income thresholds. Figure 2.1 illustrates a distribution that satisfies both Assumptions 2.2 and 2.3, while both the welfare function $W(g; P_0)$ and the budget function $B(g; P_0)$ are still continuous in g.

Figure 2.1: Illustration for Assumptions 2.1-2.3 underlying Theorem 2.1
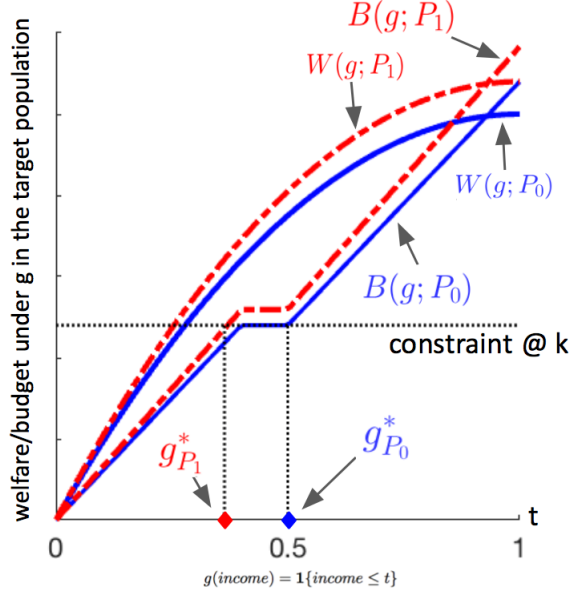
Notes: This figure plots welfare functions $W(g; P)$ and budget functions $B(g; P)$ for populations distributed according to P_0 (blue solid lines) or P_1 (red dashed lines), where P_1 is a distribution from the sequence of distributions $\{P_{h_n}\}$ that is contiguous to P_0 under Assumption 2.1. The distribution P_0 satisfies Assumptions 2.2 and 2.3. The x-axis indexes a one-dimensional criterion class $\mathcal{G} = \{g : g(x) = \mathbf{1}\{x \leq t\}\}$ for a one-dimensional continuous characteristic X_i with support on $[0, 1]$, e.g. eligibility criteria based on income thresholds. The black dotted line marks the budget threshold k. The bold blue dot marks $g_{P_0}^*$, the constrained optimal eligibility criterion under P_0. The bold red dot marks $g_{P_1}^*$, the constrained optimal eligibility criterion under P_1.

2.5 Direct extension to the existing EWM approach

The previous negative result implies that no statistical rule can be both uniformly asymptotically optimal and uniformly asymptotically feasible. Thus, policymakers might want to consider statistical rules that satisfy one of these two properties. This section demonstrates why a direct extension to the existing approach in the EWM literature does not satisfy either property.

A direct extension to the existing approach in the EWM literature is a statistical

rule that solves the sample version of the population constrained optimization problem (2):

$$\widehat{g}_s \in \arg \max_{\widehat{B}_n(g) \leq k} \widehat{W}_n(g). \tag{10}$$

The subscript s emphasizes how this approach verifies whether a criterion satisfies the constraint by comparing the *sample* analog $\widehat{B}_n(g)$ with k directly, i.e. imposes a sample-analog constraint. If no criterion satisfies the constraint, then I set $\widehat{g}_s$ to not assign any eligibility, i.e. $\widehat{g}_s(x) = 0$ for all $x \in \mathcal{X}$.

A key insight from Kitagawa and Tetenov (2018) is that without a constraint, the sample-analog rule is uniformly asymptotically optimal. Unfortunately this intuition does not extend to the current setting where the constraint involves an unknown cost. In Proposition 2.1 I find reasonable data distributions where the sample-analog rule is neither pointwise asymptotically optimal nor pointwise asymptotically feasible.
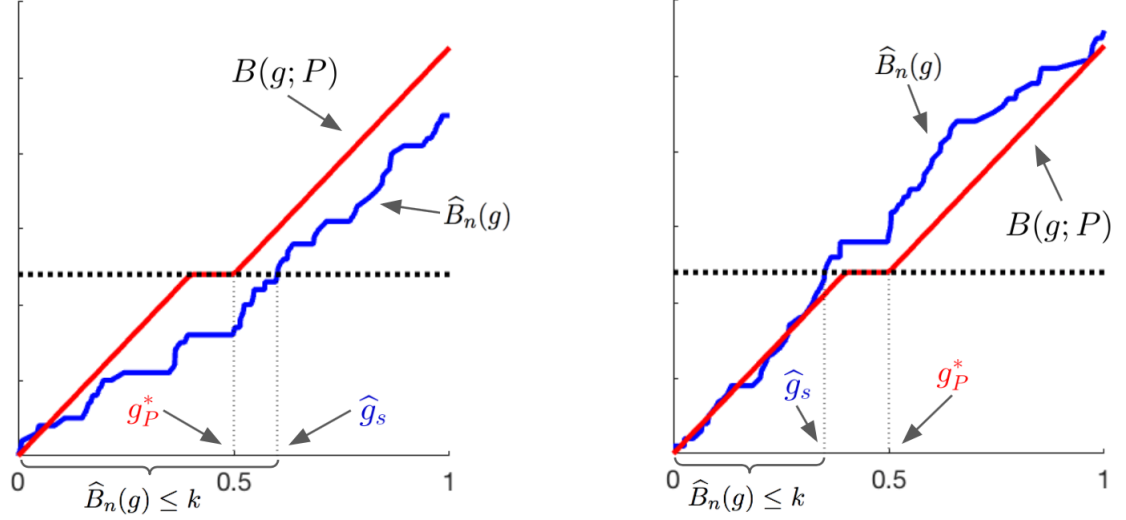
Proposition 2.1. *Consider a one-dimensional criterion class $\mathcal{G} = \{g : g(x) = \mathbf{1}\{x \leq t\}\}$, which is based on thresholds of a one-dimensional continuous characteristic X. Consider the special case where under distribution P, the benefit $\Gamma > 0$ almost surely and the cost R is binary with $R = 0$ for $X \in [\underline{t}, \overline{t}]$, but $Pr_P\{R = 1 \mid X\} \in (0, 1)$ otherwise. Then for some budget threshold $k = E_P[R \cdot \mathbf{1}\{X \leq \underline{t}\}]$, the budget constraint binds, and the sample-analog rule $\widehat{g}_s$ is neither asymptotically optimal nor asymptotically feasible under P.*

Remark 2.2. The driving force behind Proposition 2.1 is that due to sampling uncertainty, whether a criterion satisfies the sample-analog constraint is an imperfect measure of whether it satisfies the constraint in the population. Figure 2.2 illustrates the setup of Proposition 2.1, where the sampling uncertainty can be particularly problematic for the sample-analog rule.

Figure 2.2: Illustration for Proposition 2.1

(a) infeasible   (b) suboptimal

Notes: This figure plots the budget function $B(g; P)$ (red solid line) and its sample-analogs $\widehat{B}_n(g)$ (blue wriggly line) based on two different observed samples in panel (a) and (b) respectively. The x-axis indexes a one-dimensional criterion class $\mathcal{G} = \{g : g(x) = \mathbf{1}\{x \leq t\}\}$ for a one-dimensional continuous characteristic X_i with support on $[0, 1]$. The black dotted line marks the budget threshold k. The constrained optimal threshold g_P^* is $t = 0.5$. The sample-analog rule $\widehat{g}_s$ selects an infeasible threshold in panel (a) and selects a suboptimal threshold in panel (b).

Since the sample-analog rule $\widehat{g}_s$ restricts attention to criteria that satisfy the sample-analog constraint, there is no guarantee the selected criterion is actually feasible. This is very likely to happen when there is welfare gain in exceeding the budget constraint as in the setup of Proposition 2.1 where $W(g; P)$ is strictly increasing in g. Therefore, the sample-analog rule $\widehat{g}_s$ is not asymptotically feasible under P. As illustrated in Figure 2.2, after observing a sample depicted in panel (a), the sample-analog rule picks an infeasible threshold because the sample-analog constraint is still satisfied there.

Similarly, there is uncertainty about whether the sample-analog constraint is satis-

fied at the constrained optimum g_P^*, and it is possible for the sample-analog rule $\widehat{g}_s$ to miss g_P^*. In the setup of Proposition 2.1, when the sample-analog rule $\widehat{g}_s$ misses g_P^*, it is guaranteed to select a suboptimal criterion and therefore is not asymptotically optimal under P. As illustrated in Figure 2.2, after observing a sample depicted in panel (b), the sample-analog rule picks a suboptimal threshold because the sample-analog constraint is violated at the constrained optimum g_P^*. $\triangle$

Sections 3 and 4 propose statistical rules alternative to the sample-analog rule $\widehat{g}_s$, that are either uniformly asymptotically optimal or uniformly asymptotically feasible, respectively. The uniformity is considered for the class of data distributions $\mathcal{P}$ that satisfies the following assumption.

Assumption 2.4. Estimation quality. *Uniformly over $P \in \mathcal{P}$, the recentered empirical processes $\widehat{W}_n(\cdot)$ and $\widehat{B}_n(\cdot)$ defined in (6) converge to mean-zero Gaussian processes G_P^W and G_P^B uniformly over $g \in \mathcal{G}$, with covariance functions $\Sigma_P^W(\cdot,\cdot)$ and $\Sigma_P^B(\cdot,\cdot)$ respectively:*

$$\left\{ \sqrt{n} \cdot \left(\frac{1}{n} \sum_i \Gamma_i^* \cdot g(X_i) - W(g; P) \right) \right\}_{g \in \mathcal{G}} \to_d G_P^W$$

$$\left\{ \sqrt{n} \cdot \left(\frac{1}{n} \sum_i R_i^* \cdot g(X_i) - B(g; P) \right) \right\}_{g \in \mathcal{G}} \to_d G_P^B$$

The covariance functions are uniformly bounded, with diagonal entries bounded away from zero uniformly over $g \in \mathcal{G}$. There is a uniformly consistent estimator $\widehat{\Sigma}^B(\cdot,\cdot)$ of the covariance function $\Sigma_P^B(\cdot,\cdot)$.

Remark 2.3. Appendix B.3 gives primitive assumptions under which Assumption 2.4 is guaranteed to hold for $\widehat{W}_n(\cdot)$ and $\widehat{B}_n(\cdot)$ constructed using an RCT or an observational study. As standard in the literature, I need to restrict the complexity of the criterion class $\mathcal{G}$, and assume unconfoundedness and strong overlap.

3 New statistical rule that ensures uniform asymptotic feasibility

This section describes how to impose the constraint in the estimation problem (10) differently from the sample-analog rule $\widehat{g}_s$ to derive a statistical rule that is uniformly asymptotically feasible, albeit at the expense of reducing welfare. Intuitively, the statistical rule proposed in this section tightens the sample-analog constraint by taking into account estimation error. However, a tighter sample-analog constraint can lead to lower welfare if the original constraint binds at the constrained optimal rule. Theorem 3.2 formalizes this intuition.

Due to sampling uncertainty, the sample analog $\widehat{B}_n(g)$ provides an imperfect measure of the expected cost of g in the population. Thus, an infeasible criterion may be mistakenly classified as feasible by checking whether it meets the sample-analog constraint, i.e. comparing $\widehat{B}_n(g)$ directly with the constraint k. Intuitively, tightening the sample-analog constraint results in a more conservative estimate for the class of feasible eligibility criteria, and therefore reduces the chance that the selected criterion mistakenly exceeds the constraint in the population. The amount of tightening is chosen to bound the chance of a mistake uniformly over the class of distributions.

Theorem 3.1. *Suppose Assumption 2.4 holds for the class of data distributions $\mathcal{P}$. Collect eligibility criteria*

$$
\hat{\mathcal{G}}_\alpha = \left\{ g : g \in \mathcal{G} \ and \ \frac{\sqrt{n}\left(\widehat{B}_n(g) - k\right)}{\widehat{\Sigma}^B(g,g)^{1/2}} \leq c_\alpha \right\}, \tag{11}
$$

where c_α is the α-quantile from $\inf_{g \in \mathcal{G}} \left| \frac{G_P^B(g)}{\Sigma_P^B(g,g)^{1/2}} \right|$ for G_P^B the Gaussian process defined in Assumption 2.4, and $\widehat{\Sigma}^B(\cdot, \cdot)$ the consistent estimator for its covariance function. Then

$$
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} Pr_{P^n}\{\hat{\mathcal{G}}_\alpha \cap \mathcal{G}^+ \neq \emptyset\} < \alpha, \tag{12}
$$

24

where $\mathcal{G}^+ = \{g : B(g; P) > k\}$ is the set of infeasible eligibility criteria.

Note that the sample-analog constraint is tightened by $c_\alpha \cdot \frac{\widehat{\Sigma}^B (g,g)^{1/2}}{\sqrt{n}}$ where c_α is negative, which means the class $\hat{\mathcal{G}}_\alpha$ only includes eligibility criteria where the constraint is slack in the sample. The sample-analog constraint is tightened proportionally to the standard deviation to reflect the noise in $\widehat{B}_n(g)$. The degree of tightening scales inversely with the (square root of) sample size. Intuitively, larger sample size can reduce the sampling uncertainty.

3.1 Maximizing welfare while subject to the constraint probabilistically

I propose a mistake-controlling rule that maximizes the sample welfare $\widehat{W}_n(g)$ within the class of eligibility criteria $\hat{\mathcal{G}}_\alpha$ as defined in Equation (11).

$$\widehat{g}_{prob} \in \arg\max_{g \in \hat{\mathcal{G}}_\alpha} \widehat{W}_n(g). \tag{13}$$

Here the subscript *prob* highlights that with high probability this statistical rule selects feasible criterion. If the set $\hat{\mathcal{G}}_\alpha$ is empty, then I set $\widehat{g}_{prob}$ to not assign any eligibility: $\widehat{g}_{prob}(x) = 0$ for all $x \in \mathcal{X}$. Unlike $\widehat{g}_s$, as a direct consequence of Theorem 3.1, with probability at least $1 - \alpha$ this statistical rule is guaranteed to not mistakenly choose infeasible eligibility criteria. The next theorem details the improvement by $\widehat{g}_{prob}$ over $\widehat{g}_s$ in terms of uniform asymptotic feasibility.

Theorem 3.2. *Suppose Assumption 2.4 holds for the class of data distributions $\mathcal{P}$. If $\alpha_n \to 0$ as $n \to \infty$, then the mistake-controlling rule $\widehat{g}_{prob}$ defined in (13) is uniformly asymptotically feasible over $\mathcal{P}$.*

Remark 3.1. For distributions where the constraint is slack at the constrained optimal eligibility criteria, it is possible to tighten the sample-analog constraint at a rate

such that the mistake-controlling rule $\widehat{g}_{prob}$ is pointwise asymptotically optimal and pointwise asymptotically feasible. Corollary 3.1 formalizes this rate below.

Corollary 3.1. *If the data distribution P induces a constrained optimal criterion at which the constraint is slack, then the mistake-controlling rule $\widehat{g}_{prob}$ is asymptotically optimal and asymptotically feasible under P for $\alpha_n \to 0$ at a rate such that $c_{\alpha_n} = o(n^{1/2})$.*

4 New statistical rule that ensures uniform asymptotic optimality

Exceeding the budget constraint in the population might be desirable if it achieves higher welfare and does not cost too much. How to model such trade-off depends on the specific welfare program. Section 4.1 first explores possible forms of the trade-off that are of practical relevance, and then focuses on the trade-off where the marginal cost per unit of violating the constraint is constant with a known upper bound. The upper bound is context specific and as an example, Section 6 discusses how to set it for Medicaid expansion. Relaxing the constraint results in weakly higher welfare than that of the constrained optimal criterion g_P^*. Section 4.2 derives a statistical rule that implements such trade-off in the sample, and therefore is uniformly asymptotically optimal.

4.1 Form of the trade-off

Note that the original constrained optimization problem (10) may be reformulated as

$$\max_{g \in \mathcal{G}} \inf_{\lambda \geq 0} \ W(g; P) - \lambda \cdot (B(g; P) - k). \tag{14}$$

where λ measures the marginal gain of relaxing constraint. In many settings where the benefit is monetary, it may be natural to think for every dollar spent right above

the budget constraint, there is a welfare gain of λ dollar.

This formulation implies that policymakers are willing to enforce the constraint at all costs since λ is unbounded. This formulation may not reflect the real objective, however, as policymakers might only be willing to trade off violations of constraint against gains in welfare to certain extent, bounding $\lambda \in [0, \overline{\lambda}]$ and resulting in a new objective function

$$\max_{g \in \mathcal{G}} \min_{\lambda \in [0,\overline{\lambda}]} W(g; P) - \lambda \cdot (B(g; P) - k). \tag{15}$$

Its solution then balances the marginal gain and the cost from violating the constraint optimally, in an environment where the marginal gain cannot exceed $\overline{\lambda}$. This environment might be natural if the government finances the deficit by borrowing at a fixed interest rate. The trade-off can also be non-linear. Below I provide two examples. The first example allows the cost of exceeding the budget constraint to vary with the amount of deficit, perhaps because government debt can have an effect on long-term interest rates. In this setting, the violation of constraint can be modeled as a piecewise linear function:

$$\max_{g \in \mathcal{G}} \min_{\lambda_1 \in [0,\overline{\lambda}_1], \lambda_2 \in [0,\overline{\lambda}_2]} W(g; P) - \lambda_1 \cdot (B(g; P) - k_1) - \lambda_2 \cdot (B(g; P) - k_2)$$

where $k_2 > k_1$ demarcate ranges of deficit, which correspond to different interest rate λ_1 and λ_2. The second example allows the cost of exceeding the budget constraint to be constant, perhaps because violation results in a one-time penalty:

$$\max_{g \in \mathcal{G}} \min_{\lambda \in [0,\overline{\lambda}]} W(g; P) - \lambda \cdot \mathbf{1}\{B(g; P) > k\}.$$

This paper focuses on a linear trade-off as in (15). Using the notation $(x)_+$ to denote the positive part of $x \in \mathbb{R}$, the inner optimization problem of (15) can be

written as

$$\min_{\lambda \in [0, \overline{\lambda}]} W(g; P) - \lambda \cdot (B(g; P) - k) = W(g; P) - \overline{\lambda} \cdot (B(g; P) - k)_+.$$

Note, however, this reformulated objective does not involve optimization over λ: the solution to (15) is equivalent to the solution to a non-smooth but piecewise linear optimization problem:

$$\widetilde{g}_P \in \arg\max_{g \in \mathcal{G}} \; W(g; P) - \overline{\lambda} \cdot (B(g; P) - k)_+. \tag{16}$$

By construction, this reformulation is able to relax the constraint. Therefore, the solution $\widetilde{g}_P$ achieves weakly higher welfare than the constrained optimal criterion g_P^* for any data distribution P. The next lemma formalizes this observation.

Lemma 4.1. *For any $\overline{\lambda} \in [0, \infty]$, the solution to the trade-off problem (16) achieves weakly higher welfare than the constrained optimum: $W(\widetilde{g}_P; P) \geq W(g_P^*; P)$ for any distribution P.*

4.2 Implementing the trade-off in the sample

Given a new objective function (16) that trades off the gain and the cost from violating the constraint, the goal is to derive a statistical rule that is likely to select eligibility criteria that maximize the new objective function. Consider the trade-off statistical rule defined as

$$\widehat{g}_r \in \arg\max_{g \in \mathcal{G}} \; \widehat{W}_n(g) - \overline{\lambda} \cdot (\widehat{B}_n(g) - k)_+, \tag{17}$$

where the subscript r highlights that this statistical rule is able to relax the constraint.

Theorem 4.1. *Suppose Assumption 2.4 holds for the class of data distributions $\mathcal{P}$. Then the trade-off rule $\widehat{g}_r$ defined in (17) is uniformly asymptotically optimal under $\mathcal{P}$.*

To gain intuition for the above results, I note that the trade-off rule $\widehat{g}_r$ is very likely to select eligibility criteria that achieve the same welfare as $\widetilde{g}_P$ over repeated sample draws as $n \to \infty$ uniformly over $\mathcal{P}$. Furthermore, Lemma 4.1 shows the trade-off solution $\widetilde{g}_P$ achieves weakly higher welfare than the constrained optimal criterion g_P^* for any data distribution P. Therefore, the trade-off rule $\widehat{g}_r$ is asymptotically optimal uniformly over $\mathcal{P}$.

5 Simulation study

This section presents a simulation study calibrated to the distribution underlying the data from the OHIE. The simulation results confirm the negative result on the sample-analog rule $\widehat{g}_s$ discussed in Section 2.3, and illustrate the improvement by the statistical rules proposed in Sections 3 and 4.

5.1 Simulation design

To ensure the practical relevance of the simulation, I attempt to preserve the distribution of the data from the OHIE. As explained later in Section 6, I can construct benefit and cost estimates of Medicaid eligibility using the OHIE. The benefit is defined to be the increase in the probability of reporting good subjective health after receiving Medicaid eligibility, and the cost is defined to be health care expenditure that needs to be reimbursed by Medicaid, in excess to the current expansion criterion. I defer the details on the construction of the estimates (Γ_i^*, R_i^*) to Section 6.2. For the purpose of this simulation study, the OHIE represents the population P, and therefore I can take these estimates as the true benefit and cost (Γ, R).

Table 5.1 presents the basic summary statistics for the estimates (Γ_i^*, R_i^*). Under the current definition of the cost, a criterion is feasible if it incurs a negative cost.

Table 5.1: Summary statistics of the OHIE sample by number of children

Number of children	Sample size	Sample mean of Γ_i^*	Sample mean of R_i^*
0	5,758	3.1%	$651
1	1,736	10.3%	$348
≥ 2	2,641	1.5%	-$275
-	10,135	3.9%	$358

Notes: This table presents summary statistics on the sample of individuals who responded to both the initial and the main surveys from the Oregon Health Insurance Experiment (the OHIE sample). The first three rows represent individuals living with different number of children (family members under age 19), and the last row is the aggregate. The estimate for benefit Γ_i^* is an estimate for the increase in the probability of an individual reporting "excellent/very good/good" on self-reported health (as opposed to "poor/fair") after receiving Medicaid eligibility. The estimate for cost R_i^* is an estimate for individual's health care expenditure that needs to be reimbursed by Medicaid, in excess of the current expansion criteria.

I randomly draw observations from the OHIE sample to form a random sample. Under this simulation design, I can solve for the constrained optimal criterion as

$$g_P^* \in \arg \max_{g \in \mathcal{G}, \ B(g;P) \leq 0} W(g; P)$$

where $(W(g; P), B(g; P))$ are the sample-analog in the OHIE sample. The criterion class $\mathcal{G}$ includes income thresholds that can vary with the number of children. The constrained optimal criterion is

$$g_P^*(x) = \begin{cases} \mathbf{1}\{income \leq 75\%\}, & numchild = 0 \\ \mathbf{1}\{income \leq 343\%\}, & numchild = 1 \\ \mathbf{1}\{income \leq 160\%\}, & numchild \geq 2 \end{cases}$$

for characteristics $x = (income, \ numchild)$. The maximum feasible welfare is given by $W(g_P^*; P) = 3.8\%$, an increase of 3.8% in reporting good subjective health. The cost associated with the constrained optimal criterion is $B(g_P^*; P) = -\$0.6$, meaning

per enrollee g_P^* costs \$0.6 less than the current expansion criterion.

5.2 Simulation results

Table 5.2 compares the performance of various statistical rules $\widehat{g}$ through 500 Monte Carlo iterations. At each iteration, I randomly draw observations from the OHIE sample to form a random sample. I simulate with the same sample size as the original sample to hold the amount of sampling uncertainty constant. Given the random sample, I collect eligibility criteria chosen by each of following statistical rules:

- sample-analog rule $\widehat{g}_s$.

- mistake-controlling rule $\widehat{g}_{prob}$ with critical value c_α described in Section 3 (α is set to 5%.)

- trade-off rule $\widehat{g}_r$ with trade-off coefficient $\bar{\lambda}$ described in Section 6.3.2.

I evaluate the welfare function and the budget function $(W(g; P), B(g; P))$ for a given criterion in the original OHIE sample. Averages over 500 iterations provide simulation evidence on the asymptotic properties of the above statistical rules, as shown in Table 5.2.

Table 5.2: Simulation results: asymptotic properties of statistical rules $\widehat{g}$

Statistical rule	sample-analog $\widehat{g}_s$	mistake-controlling $\widehat{g}_{prob}$	trade-off $\widehat{g}_r$
Prob. of selecting infeasible criteria	35.4%	8%	79.6%
Prob. of selecting suboptimal criteria	87.0%	98.6%	37.6%
Average welfare loss	0.06	0.60	-0.02
Average cost	-\$3	-\$57	\$105

Notes: This table reports asymptotic properties of statistical rules $\widehat{g}$, as averaged over 500 simulations. Row 1 reports the probability that the rule selects an eligibility criterion that violates the budget constraint, i.e. $Pr_{P^n}\{B(\widehat{g}; P) > 0\}$. Row 2 reports the probability that the rule achieves strictly less welfare than the constrained optimal criterion g_P^*, i.e. $Pr_{P^n}\{W(\widehat{g}; P) < W(g_P^*; P)\}$. Row 3 reports the average welfare loss of the rule relative to the maximum feasible welfare, i.e. $\frac{E_{P^n}\left[W(g_P^*;P) - W(\widehat{g};P)\right]}{W(g_P^*;P)}$. Row 4 reports the average cost of the criteria selected by the rule, i.e. $E_{P^n}\left[B(\widehat{g}; P)\right]$.

Row 1 of Table 5.2 illustrates that it is possible for all three statistical rule $\widehat{g}$ to select infeasible criteria. A lower probability of selecting infeasible criteria suggests the rule is closer to achieving asymptotic feasibility. Proposition 2.1 describes distributions where the sample-analog rule $\widehat{g}_s$ is not asymptotic feasible. In the distribution calibrated to the OHIE sample, the sample-analog rule $\widehat{g}_s$ might not be asymptotically feasible, either, as it can select infeasible eligibility criteria in 19.4% of the draws. In contrast, Theorem 3.1 guarantees that the mistake-controlling rule $\widehat{g}_{prob}$ selects infeasible eligibility criteria in less than 5% of the draws, regardless of the distribution. Simulation confirms such guarantee as the mistakes only happen 8% of the time.

Row 2 of Table 5.2 illustrates that it is possible for all three statistical rule $\widehat{g}$ to achieve weakly higher welfare than the constrained optimal criterion g_P^*. This can happen when $\widehat{g}$ selects an infeasible criterion. A lower probability of selecting suboptimal criteria suggests the rule is closer to achieving asymptotic optimality. Theorem 4.1 implies that the trade-off rule $\widehat{g}_r$ is uniformly asymptotically optimal while there is no such guarantee for the sample-analog rule $\widehat{g}_s$.

In the distribution calibrated to the OHIE, the trade-off rule $\widehat{g}_r$ on average achieves higher welfare than the sample-analog rule $\widehat{g}_s$. As shown in row 3 of Table 5.2, the welfare loss of $\widehat{g}_r$ is -2% of the maximum feasible welfare $W(g_P^*; P)$, compared to 6% for $\widehat{g}_s$. However, its improvement can be at the cost of violating the budget constraint more often than $\widehat{g}_s$, at a rate of 79.6%. Comparing the distribution of the cost $B(\widehat{g}_s; P)$ and $B(\widehat{g}_r; P)$, I note whenever the trade-off rule $\widehat{g}_r$ selects an infeasible criterion, the amount of violation is small, so that on average the budget constraint would not be violated as shown in row 4 of Table 5.2. As a result, even though the trade-off rule $\widehat{g}_r$ is more likely to select infeasible eligibility criteria than the sample-analog rule $\widehat{g}_s$, the cost to these violations is limited.

6 Empirical example

I now apply my methods to improve the current Medicaid expansion criterion. Section 6.1 includes background information on the current Medicaid expansion criterion and explains the need for a more flexible one. Section 6.2 overviews the Oregon Medicaid Health Insurance Experiment (OHIE). Section 6.3 describes how to implement the two methods I propose to select a more flexible criterion using results from the OHIE. Section 6.4 provides some practical guidance on how to choose between these two methods.

6.1 Medicaid expansion

Medicaid is a government-sponsored health insurance program intended for the low-income population in the United States. Up till 2011, many states provided Medicaid eligibility to able-bodied adults with income up to 100% of the federal poverty level. The 2011 Affordable Care Act (ACA) provided resources for states to expand Medicaid eligibility for all adults with income up to 138% of the federal poverty level starting in 2014. The current expansion criterion in Oregon is therefore to provide

Medicaid eligibility to everyone with income below 138% of the federal poverty level, regardless of other characteristics. In this empirical example, I want to maximize the benefit of Medicaid eligibility by allowing a more flexible expansion criterion than the current one. Specifically, the more flexible criterion would allow the income thresholds to vary with other characteristics, such as number of children.

While the goal of the expansion is to insure more low-income individuals, given eligibility, enrollment in Medicaid is not mandatory because eligible individuals can choose not to enroll. Therefore I define Γ as the benefit of Medicaid eligibility instead of the benefit of Medicaid enrollment. Finkelstein et al. (2019) model how eligible individuals value Medicaid in terms of both the health benefit and their ability to smooth consumption from having health insurance. For illustration I focus on the health benefit, and define the benefit as to the increase in the probability of reporting good subjective health after becoming eligible for Medicaid. Alternatively, focusing on the benefit in terms of subjective health can be justified by the health-based optimization approach as in Finkelstein et al. (2019).

It may be infeasible to implement a new expansion criterion if it costs more than the current one. Therefore I impose the constraint that per enrollee, the more flexible criterion must cost no more than the current one. A way to model the cost of the new expansion criterion is to assume that everyone made eligible for Medicaid would enroll. The resulting cost measure is conservative because in reality, not everyone eligible for Medicaid would eventually enroll. Since an eligible individual does not contribute to government's spending if she decides not to enroll, I can tighten the conservative cost measure by taking into account her enrollment decision. The corresponding constraint thus involves both individual's enrollment decision and health care utilization after becoming eligible for Medicaid.

In brief, the input to the constrained optimization is $A = (\Gamma, R, X)$. While both Γ and R are unknown quantities, the randomization of the Medicaid eligibility from the OHIE provides good estimates for them, as explained below.

6.2 Data

I use the experimental data from the OHIE, where Medicaid eligibility (D_i) was randomized in 2007 among Oregon residents who were low-income adults, but previously ineligible for Medicaid, and who expressed interest in participating in the experiment. Finkelstein et al. (2012) include a detailed description of the experiment and an assessment of the average effects of Medicaid on health and health care utilization. I include a cursory explanation here for completeness, emphasizing that my final sample differs from the original OHIE sample as I focus on individuals who responded both to the initial and the main surveys from the OHIE.

The original OHIE sample consists of 74,922 individuals (representing 66,385 households). Of these, 26,423 individuals responded to the initial mail survey, which collects information on income as percentage of the federal poverty level and number of children.[6] Recall that I want to consider a more flexible expansion criterion that allows the income threshold to vary with number of children. Therefore, the initial survey include the characteristics X of interest, and I exclude individuals who did not respond to the initial survey from my sample.

Due to this difference, the expansion criteria selected using my sample do not directly carry their properties to the population underlying the original OHIE sample, as the distributions of X differ.

The main survey collects data related to health (Y_i), health care utilization (C_i) and actual enrollment in Medicaid (M_i), which allows me to construct estimates for the other two inputs, namely (Γ, R). Therefore I further exclude individuals who did not respond to the main survey from my sample. More details on the main survey data and the construction of these estimates follow.

For health (Y_i), I follow the binary measurement in Finkelstein et al. (2012) based on self-reported health, where an answer of "poor/fair" is coded as $Y_i = 0$ and "ex-

---

[6]More accurately, I follow Sacarny et al. (2020) to approximate number of children by the number of family members under age 19 living in house as reported on the initial mail survey.

cellent/very good/good" is coded as $Y_i = 1$. For health care utilization (C_i), the study collected measures of utilization of prescription drugs, outpatient visits, ER visits, and inpatient hospital visits. Finkelstein et al. (2012) annualize these utilization measures to turn these into spending estimates, weighting each type by its average cost (expenditures) among low-income publicly insured non-elderly adults in the Medical Expenditure Survey (MEPS). Note that health and health care utilization are not measured at the same scale, which requires rescaling when I consider the trade-off between the two. I address this issue in Section 6.3. Lastly, since the enrollment in Medicaid still requires an application, not everyone eligible in the OHIE eventually enrolled in Medicaid, which implies $M_i \leq D_i$.

Given the setup of the OHIE, Medicaid eligibility (D_i) is random conditional on household size (number of adults in the household) entered on the lottery sign-up form and survey wave. While the original experimental setup would ensure randomization given household size, the OHIE had to adjust randomization for later waves of survey respondents (see the Appendix of Finkelstein et al. (2012) for more details). Denote the confounders (household size and survey wave) with V_i, and define the propensity score as $p(V_i) = Pr\{D_i = 1 \mid V_i\}$. If the propensity score is known, then the construction of the estimates follows directly from the formula (7) as in Section 2.2. However, the adjustment for later survey waves means I need to estimate the propensity score, and I adapt the formula (7) following Athey and Wager (2020) to account for the estimated propensity score.

Specifically, define the conditional expectation function (CEF) of a random variable U_i as $\gamma^U = E[U_i \mid V_i, D_i]$. Since V_i in my case is discrete, I use a fully saturated model to estimate the propensity score $\widehat{p}(V_i)$ and the CEF $\widehat{\gamma}^U(V_i, D_i)$. I then form the estimated Horvitz-Thompson weight with the estimated propensity score as $\widehat{\alpha}(V_i, D_i) = \frac{D_i}{\widehat{p}(V_i)} - \frac{1-D_i}{1-\widehat{p}(V_i)}$. Define $\Gamma_i^* = \widehat{\gamma}^Y(V_i, 1) - \widehat{\gamma}^Y(V_i, 0) + \widehat{\alpha}(V_i, D_i) \cdot \left(Y_i - \widehat{\gamma}^Y(V_i, D_i)\right)$. Then as shown in the Appendix B.3, the estimate Γ_i^* is a noisy version of Γ, the health benefit due to Medicaid eligibility, and the noise is asymp-

totically unimportant.

The 2014 Medicaid spending was roughly \$6,000 per adult enrollee in Oregon, according to the expenditure information obtained from MACPAC (2019). To formalize the budget constraint that the per enrollee cost of the proposed criterion cannot exceed the 2014 criterion, I need to account for imperfect take-up because not everyone eligible for Medicaid would enroll. Define $R_i^* = \widehat{\gamma}^Z(V_i, 1) + \frac{D_i}{\widehat{p}(W_i)} \cdot \left(Z_i - \widehat{\gamma}^Z(V_i, D_i) \right)$ where $Z_i = C_i - \$6,000 \cdot M_i$. Then as shown in the Appendix B.3, the estimate R_i^* is a noisy version of R, the per enrollee *excess* cost relative to the current level, and the noise is asymptotically unimportant.

6.3 Budget-constrained optimal Medicaid expansion

This section reports the Medicaid expansion criteria selected by the two statistical rules proposed in Sections 3 and 4 using results from the OHIE. The following subsections provide details on implementing these statistical rules.

Figure 6.1 summarizes the selected expansion criteria, which are income thresholds specific to the number of children. The mistake-controlling rule $\widehat{g}_{prob}$ chooses to restrict Medicaid eligibility, especially lowering the income threshold for childless individuals far below the current level. This reflects the large variation in the cost estimates, which results in uncertainty about whether it would be feasible to provide eligibility to many individuals in the population. As a result, the budget estimate for the selected criterion is -\$140, far below the threshold of zero. In contrast, the trade-off rule $\widehat{g}_r$ chooses to assign Medicaid eligibility to more individuals, and to raise the income thresholds above the current level. The higher level occurs because on average the benefit estimates are positive, as illustrated in Table 5.1, which suggests many individuals still exhibit health benefit from being eligible for Medicaid. Under a reasonable assumption for an upper bound on the monetary value on the health benefit specified in Section 6.3.2, the trade-off rule finds that the additional health benefit from violating the budget constraint exceeds the cost of doing so. Therefore

the budget estimate for the selected criterion is \$110, slightly above the threshold of zero.

Figure 6.1: More flexible Medicaid expansion criteria

(a) criterion selected by the mistake-controlling rule

(b) criterion selected by the trade-off rule
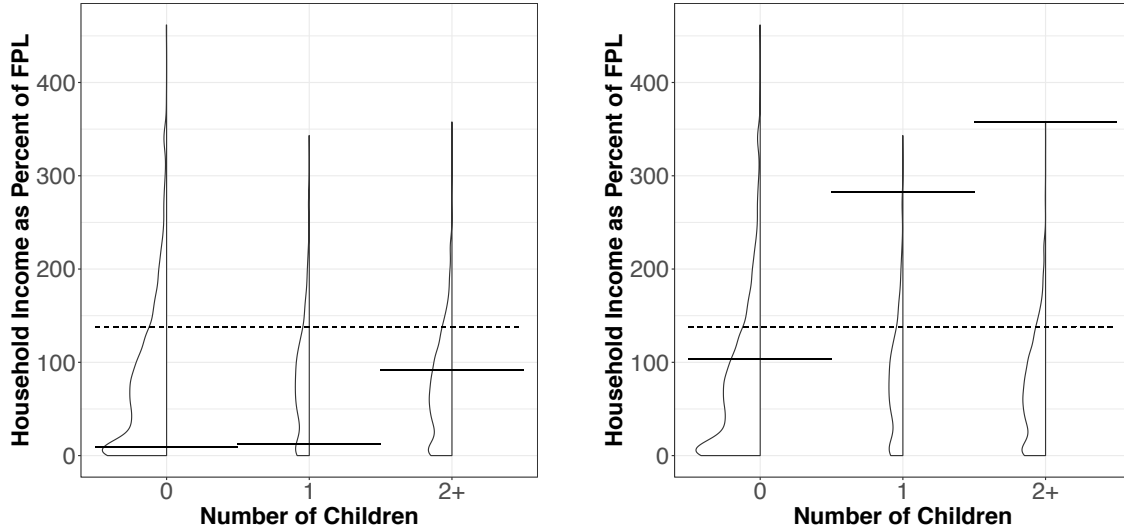
Notes: This figure plots the more flexible Medicaid expansion criteria selected by the two statistical rules proposed in Sections 3 and 4 based on results from the OHIE. The horizontal dashed line marks the income thresholds under the current expansion criterion, which is 138% regardless of the number of children in a household. The horizontal solid lines mark the more flexible criterion selected by various statistical rules, i.e. income thresholds that can vary with number of children. For each number of children, I also plot the underlying income distribution to visualize individuals below the thresholds. Panel (a) plots the criterion selected by the mistake-controlling rule $\widehat{g}_{prob}$. Panel (b) plots the criterion selected by the trade-off rule $\widehat{g}_r$.

6.3.1 Implementing the mistake-controlling rule

To construct the mistake-controlling rule $\widehat{g}_{prob}$ as proposed in Section 3, I maximize the sample welfare function among eligibility criteria in $\hat{\mathcal{G}}_\alpha$, which are guaranteed to contain only feasible eligibility criteria with probability approaching $1 - \alpha$. For a conventional level of $\alpha = 5\%$, constructing $\hat{\mathcal{G}}_\alpha$ requires an estimate for the critical value c_α, the 5%-quantile from $\inf_{g \in \mathcal{G}} \frac{G_P^B(g)}{\Sigma_P^B(g,g)^{1/2}}$, the infimum of the Gaussian process

38

$G_P^B(\cdot) \sim \mathcal{GP}(0, \Sigma_P^B(\cdot, \cdot))$. In practice, I construct a grid on $\mathcal{G}$ as

$$\widetilde{\mathcal{G}} = \{g(x) = \mathbf{1}\{income \cdot \mathbf{1}\{numchild = j\} \leq y_j\} : j \in \{0, 1, \geq 2\}, y_j \in \{0, 50, 100, \ldots, 500\}\} \tag{18}$$

for characteristics $x = (income, \; numchild)$. This grid thus consists of income thresholds every 50% of the federal poverty level, and the thresholds can vary with number of children. I then approximate the infimum over infinite-dimensional $\mathcal{G}$ by the minimum over $\widetilde{\mathcal{G}}$ with estimated covariance i.e. $\min\limits_{g \in \widetilde{\mathcal{G}}} \frac{\widetilde{h}(g)}{\widehat{\Sigma}^B(g,g)^{1/2}}$. Here $\widetilde{h}(\cdot) \sim \mathcal{N}(0, \widehat{\Sigma}^B)$ is a Gaussian vector indexed by $g \in \widetilde{\mathcal{G}}$, with $\widehat{\Sigma}^B$ is sub-matrix of the covariance estimate $\widehat{\Sigma}^B(\cdot, \cdot)$ for $g \in \widetilde{\mathcal{G}}$. Based on 10,000 simulation draws I estimate c_α to be -2.56. The validity of this approximation is given by the uniform consistency of the covariance estimator under Assumption 2.4.

6.3.2 Implementing the trade-off rule

To construct the trade-off rule $\widehat{g}_{prob}$ as proposed in Section 4, I need to choose $\overline{\lambda}$, the upper bound on the marginal gain from violating the constraint. In my empirical illustration, the budget constraint is in terms of monetary value. The objective function, however, is measured based on self-reported health, which does not directly translate to a monetary value. Following Finkelstein et al. (2019), I convert self-reported health into value of a statistical life year (VSLY) based on existing estimates. Specifically, a conservative measure for the increase in quality-adjusted life year (QALY) when self-reported health increases from "poor/fair" to "excellent/very good/good" is roughly 0.6. The "consensus" estimate for the VSLY for one unit of QALY from Cutler (2004) is $100,000 for the general US population. Taken these estimates together, I set $\overline{\lambda} = 1/(0.6 \cdot 10^5)$.

6.4 Practical guidance

Table 6.1 summarizes the theoretical results and compares the properties for each of the statistical rule. Importantly, the obvious extension to the existing approach $\widehat{g}_s$ satisfies neither of the desirable properties, and the proposed rules $\widehat{g}_{prob}$ and $\widehat{g}_r$ focus on different properties. A natural question is how policymakers should choose the rules. I argue below this choice should align with policymakers' attitude toward the budget constraint.

Table 6.1: Summary of properties for each rule

Statistical rule	sample-analog $\widehat{g}_s$	mistake-controlling $\widehat{g}_{prob}$	trade-off $\widehat{g}_r$
Uniform asymptotic optimality	$\times$	$\times$	$\checkmark$
Uniform asymptotic feasibility	$\times$	$\checkmark$	$\times$

The mistake-controlling rule $\widehat{g}_{prob}$ selects feasible eligibility criteria with high probability while trying to achieve the highest benefit possible. If policymakers are financially conservative, then eligibility criteria selected by $\widehat{g}_{prob}$ are more suitable. The trade-off rule $\widehat{g}_r$ selects eligibility criteria that achieve at least the maximized feasible welfare. When $\widehat{g}_r$ achieves even higher welfare than the maximized feasible welfare, $\widehat{g}_r$ selects an infeasible eligibility criterion but accounts for the per unit return from violating the constraint. If policymakers want to reach as many individuals as they can, then eligibility criteria selected by $\widehat{g}_r$ are more suitable. In the particular distribution of the OHIE sample, the simulation study in Section 5 suggests that the cost of these violations may not be prohibitively high. Since the exact amount of the trade-off depends on the joint distribution of the benefit and cost (Γ, R), policymakers need to assess the amount of the trade-off for their particular settings.

7 Conclusion

In this paper, I extend the existing EWM approach to allow for a budget constraint where the cost of implementing any given eligibility criterion needs to be estimated. The existing EWM approach directly maximizes a sample analog of the social welfare function, and only accounts for constraints that can be verified with certainty in the population. In reality, the cost of providing eligibility to any given individual might be unknown ex-ante due to imperfect take-up and heterogeneity. Therefore, one cannot verify with certainty whether any given eligibility criterion satisfies the budget constraint in the population. I propose two alternative statistical rules: the mistake-controlling rule and the trade-off rule. I illustrate their asymptotic properties and implementation details using experimental data from the OHIE.

Each of these two rules satisfies one of the two desirable properties: uniform asymptotic feasibility and uniform asymptotic optimality, where uniformity is imposed on a broad class of data distributions. A promising avenue for future research is whether within the class of uniformly asymptotic feasible rules, the mistake-controlling rule is the closest to achieving uniform asymptotic optimality. Likewise, it remains to be seen whether the the trade-off rule is asymptotical feasibility for a specific class of data distributions.

References

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks," *ProPublica*, May 2016.

Athey, Susan and Stefan Wager, "Policy Learning with Observational Data," *Econometrica*, 2020, (Forthcoming).

Bhattacharya, Debopam and Pascaline Dupas, "Inferring welfare maximizing treatment assignment under budget constraints," *Journal of Econometrics*, March 2012, *167* (1), 168–196.

Chamberlain, Gary, "Bayesian Aspects of Treatment Choice," in "The Oxford Handbook of Bayesian Econometrics," New York: Oxford University Press, September 2011, pp. 11–39.

Cutler, David, *Your Money or Your Life: Strong Medicine for America's Health Care System*, New York, New York: Oxford University Press, 2004.

Dehejia, Rajeev H., "Program evaluation as a decision problem," *Journal of Econometrics*, 2005, *125* (1-2), 141–173.

Demirer, Mert, Vasilis Syrgkanis, Greg Lewis, and Victor Chernozhukov, "Semi-Parametric Efficient Policy Learning with Continuous Actions," *arXiv:1905.10116 [cs, econ, math, stat]*, July 2019.

Dhailiwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch, "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education," in "Education Policy in Developing Countries," University of Chicago Press, December 2013.

Finkelstein, Amy and Matthew J. Notowidigdo, "Take-Up and Targeting: Experimental Evidence from SNAP," *The Quarterly Journal of Economics*, August 2019, *134* (3), 1505–1556.

_ , **Matthew Gentzkow, Peter Hull, and Heidi Williams**, "Adjusting Risk Adjustment — Accounting for Variation in Diagnostic Intensity," *New England Journal of Medicine*, February 2017, *376* (7), 608–610.

_ , **Nathaniel Hendren, and Erzo F. P. Luttmer**, "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment," *Journal of Political Economy*, December 2019, *127* (6), 2836–2874.

_ , **Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker**, "The Oregon Health Insurance Experiment: Evidence from the First Year," *The Quarterly Journal of Economics*, August 2012, *127* (3), 1057–1106.

Gelber, Alexander, Adam Isen, and Judd B. Kessler, "The Effects of Youth Employment: Evidence from New York City Lotteries," *The Quarterly Journal of Economics*, February 2016, *131* (1), 423–460.

Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack, "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago," *The Quarterly Journal of Economics*, February 2017, *132* (1), 1–54.

Hendren, Nathaniel and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 2020, *135* (3), 1209–1318.

Hirano, Keisuke and Jack R. Porter, "Asymptotics for Statistical Treatment Rules," *Econometrica*, 2009, *77* (5), 1683–1701.

Kasy, Maximilian and Anja Sautmann, "Adaptive Treatment Assignment in Experiments for Policy Choice," *Econometrica*, 2020, (Forthcoming).

Kearns, Michael, Seth Neel, Aaron Roth, and Zhiwei Steven Wu, "Preventing Fairness Gerrymandering: Auditing and Learning for Subgroup Fairness," in "International Conference on Machine Learning," Vol. 80 of *Proceedings of Machine Learning Research* July 2018, pp. 2564–2572.

Kitagawa, Toru and Aleksey Tetenov, "Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice," *Econometrica*, 2018, *86* (2), 591–616.

MACPAC, "EXHIBIT 22. Medicaid Benefit Spending Per Full-Year Equivalent (FYE) Enrollee by State and Eligibility Group," 2019.

Manski, Charles F., "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 2004, *72* (4), 1221–1246.

Mbakop, Eric and Max Tabord-Meehan, "Model Selection for Treatment Choice: Penalized Welfare Maximization," *arXiv:1609.03167 [econ, math, stat]*, December 2019.

Rai, Yoshiyasu, "Statistical Inference for Treatment Assignment Policies," 2019.

Sacarny, Adam, Katherine Baicker, and Amy Finkelstein, "Out of the Woodwork: Enrollment Spillovers in the Oregon Health Insurance Experiment," Working Paper 26871, National Bureau of Economic Research March 2020.

Stoye, Jörg, "Minimax regret treatment choice with finite samples," *Journal of Econometrics*, 2009, *151* (1), 70–81.

Sun, Hao, Shuyang Du, and Stefan Wager, "Treatment Allocation under Uncertain Costs," *arXiv:2103.11066 [stat]*, March 2021. arXiv: 2103.11066.

van der Vaart, A. W., "Asymptotic Statistics by A. W. van der Vaart," October 1998.

van der Vaart, Aad W. and Jon A. Wellner, *Weak Convergence and Empirical Processes* Springer Series in Statistics, New York: Springer, 1996.

Wainwright, Martin J., "High-Dimensional Statistics by Martin J. Wainwright," February 2019.

Online Appendix

A Proofs of theorems

Proof. **Proof of Theorem 2.1.** The population is a probability space $(\Omega, \mathcal{A}, P)$, which induces the sampling distribution P^n that governs the observed sample. A statistical rule $\widehat{g}$ is a mapping $\widehat{g}(\cdot) : \Omega \to \mathcal{G}$ that selects a policy from the policy class $\mathcal{G}$ based on the observed sample. Note that the selected policy $\widehat{g}(\omega)$ is still deterministic because the policy class $\mathcal{G}$ is restricted to be deterministic policies, though the proof below extends to random policies. When no confusion arises, I drop the reference to event ω for notational simplicity.

Suppose $\widehat{g}$ is asymptotically optimal and asymptotically feasible under P_0. We want to prove there is non-vanishing chance that $\widehat{g}$ selects policies that are infeasible some other distributions:

$$\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} Pr_{P^n} \left\{ \omega : B(\widehat{g}(\omega); P) > k \right\} > 0. \tag{19}$$

Asymptotical optimality under P_0 implies for any $\epsilon > 0$ we have

$$\limsup_{n \to \infty} Pr_{P_0^n} \left\{ \omega : W(g_{p_0}^* P_0) - W(\widehat{g}(\omega); P_0) > \epsilon \right\} = 0. \tag{20}$$

Asymptotical feasibility under P_0 implies $Pr_{P_0^n}\{\omega : B(\widehat{g}(\omega); P_0) \leq k\} \to 1$.

Consider the event ω' where $|W(\widehat{g}(\omega'); P) - W(g_P^*; P)| < \epsilon$ and $\widehat{g}(\omega')$ is feasible. Asymptotic optimality and asymptotic feasibility imply imply $Pr_{P_0^n} \{\omega'\} \to 1$. To see this, note the probability of such event has an asymptotic lower bound of one

$$Pr_{P_0^n} \left\{ \omega' : W(g_{p_0}^* P_0) - W(\widehat{g}(\omega'); P_0) \leq \epsilon \text{ and } B(\widehat{g}(\omega'); P_0) \leq k \right\}$$
$$\geq Pr_{P_0^n} \left\{ \omega : W(g_{p_0}^* P_0) - W(\widehat{g}(\omega); P_0) \leq \epsilon \right\} + Pr_{P_0^n} \left\{ \omega : B(\widehat{g}(\omega); P_0) \leq k \right\} - 1$$

where the first two terms converge to one as $n \to \infty$ respectively under asymptotic optimality and asymptotic feasibility. For ϵ satisfying Assumption 2.3, we have $B(\widehat{g}; P_0) = B(g^*_{P_0}; P_0) = k$ under the event ω', since the constraint is exactly satisfied at $g^*_{P_0}$ under Assumption 2.2. By Law of Total Probability, we have

$$Pr_{P_0^n} \{\omega : B(\widehat{g}(\omega); P_0) = k\} \geq Pr_{P_0^n} \{\omega'\} \cdot Pr_{P_0^n} \{B(\widehat{g}; P_0) = k \mid \omega'\}$$

Then the above argument shows $Pr_{P_0^n} \{\omega : B(\widehat{g}(\omega); P_0) = k\} \to 1$ as $n \to \infty$.

Following the notation in Assumption 2.1, denote the set of policies where the constraints bind exactly under the limit distribution P_0 by

$$\mathcal{G}_0 = \{g \in \mathcal{G} : B(g; P_0) = k\}. \tag{21}$$

Under Assumption 2.1, the sequence $P_{h_n}^n$ is contiguous with respect to the sequence P_0^n, which means $P_0^n(A_n) \to 0$ implies $P_{h_n}^n(A_n) \to 0$ for every sequence of measurable sets A_n on $\mathcal{A}^n$. Then $Pr_{P_0^n} \{A_n : B(\widehat{g}(A_n); P_0) = k\} \to 1$ implies there exists an $N(u)$ such that for all $n \geq N(u)$, we have $Pr_{P_{h_n}^n} \{A_n : B(\widehat{g}(A_n); P_0) = k\} \geq 1 - u$. That is, with high probability, the statistical rule $\widehat{g}$ selects policies from $\mathcal{G}_0$ based on the observed sample distributed according to $P_{h_n}^n$. Recall Assumption 2.1 implies for all $g \in \mathcal{G}_0$, for any sample size n, we have $B(g; P_{h_n}) - k > C/\sqrt{n}$. Thus this statistical rule cannot uniformly satisfy the constraint since with sample size $n \geq N(u)$, we have

$$\sup_{P \in \mathcal{P}} Pr_{P^n} \{B(\widehat{g}; P) > k\} \geq Pr_{P_{h_n}^n} \{B(\widehat{g}; P_{h_n}) > k\} \geq 1 - u \tag{22}$$

$\square$

Proof. **Proof of Proposition 2.1.** Consider a one-dimensional policy class $\mathcal{G} = \{g : g(x) = \mathbf{1}\{x \leq t\}\}$, which includes thresholds for a one-dimensional continuous characteristic X. Since $\Gamma > 0$ almost surely, the welfare function $W(t; P) := E_P[\Gamma \cdot \mathbf{1}\{X \leq t\}]$ is strictly increasing in t.

Suppose the budget constraint takes the form of a capacity constraint, involving a binary take-up decision R (that may or may not be independent of X). By assumption, the probability an individual takes up the treatment for $X \in [\underline{t}, \overline{t}]$ is zero but between zero and one otherwise. Then the budget function $B(t; P) := E_P[R \cdot \mathbf{1}\{X \leq t\}]$ is flat in the interval $[\underline{t}, \overline{t}]$ but also strictly increasing otherwise.

The population problem is

$$\max_t \ W(t; P) \ s.t. \ B(t; P) \leq k,$$

where $B(t; P) = k$ for $t \in [\underline{t}, \overline{t}]$ by assumption. The constrained optimal threshold is therefore the highest threshold where the constraint is satisfied exactly i.e. $t^* = \overline{t}$. This also implies $R \cdot \mathbf{1}\{X \leq t\} \sim Bernoulli(k)$ for $t \in [\underline{t}, \overline{t}]$.

The sample-analog rule solves the following sample problem

$$\max_t \frac{1}{n} \sum_i \Gamma_i \cdot \mathbf{1}\{X_i \leq t\} \ s.t. \ \widehat{B}_n(t) \leq k$$

for $\widehat{B}_n(t) := \frac{1}{n} \sum_i R_i \cdot \mathbf{1}\{X_i \leq t\}$. Given that $\Gamma > 0$ almost surely, the sample-analog rule equivalently solves $\max_t \widehat{B}_n(t) \ s.t. \ \widehat{B}_n(t) \leq k$. However, the solution is not unique because $\widehat{B}_n(t)$ is a step function. To be conservative, let the sample-analog rule be the smallest possible threshold to maximize $\widehat{B}_n(t)$:

$$\widehat{t} = \min \left\{ \arg\max_t \{\widehat{B}_n(t) \ s.t. \ \widehat{B}_n(t) \leq k\} \right\}.$$

Note that we can also write $\widehat{B}_n(t) = \frac{1}{n} \sum_{R_i=1} \mathbf{1}\{X_i \leq t\}$, which makes it clear that $\widehat{t}$ corresponds to ranking X_i among individuals with $R_i = 1$, and then picking the lowest

threshold such that we assign treatment to the first $\lfloor k \cdot n \rfloor$ individuals. This also means if in the sample few individuals take up the treatment such that $\frac{1}{n} \sum_i R_i \leq k$, we can have a sample-analog rule that treats everyone up to $\max_{R_i=1} X_i$. Taken together both scenarios, we note the sample-analog rule implies the treated share in the sample is equal to

$$\widehat{B}_n(\widehat{t}) := \frac{1}{n} \sum_i R_i \cdot \mathbf{1}\{X_i \leq \widehat{t}\} = \min\left\{\frac{1}{n} \sum_i R_i, \frac{\lfloor k \cdot n \rfloor}{n}\right\}.$$

Note that
$$\begin{cases} \widehat{t} > \overline{t} \Leftrightarrow & B(\widehat{t}; P) > B(\overline{t}; P) \\ \widehat{t} < \underline{t} \Leftrightarrow & W(\widehat{t}; P) < W(\underline{t}; P) \end{cases}$$

which means whenever $\widehat{t} > \overline{t}$, the sample-analog rule violates the constraint in the population as $B(\overline{t}; P) = k$; whenever $\widehat{t} < \underline{t}$, the sample-analog rule achieves strictly less welfare than t^* in the population because $W(\underline{t}; P)$ is strictly less than $W(t^*; P)$. We next derive the limit probability for these two events. Applying Law of Total Probability, we have

$$Pr_{P^n}\left\{\widehat{B}_n(\overline{t}) < \widehat{B}_n(\widehat{t})\right\}$$

$$= Pr_{P^n}\left\{\widehat{B}_n(\overline{t}) < \frac{\lfloor k \cdot n \rfloor}{n} \text{ and } \frac{\lfloor k \cdot n \rfloor}{n} < \frac{1}{n} \sum_i R_i\right\}$$

$$+ Pr_{P^n}\left\{\widehat{B}_n(\overline{t}) < \frac{1}{n} \sum_i R_i \text{ and } \frac{\lfloor k \cdot n \rfloor}{n} \geq \frac{1}{n} \sum_i R_i\right\} \qquad (23)$$

$$\geq Pr_{P^n}\left\{\widehat{B}_n(\overline{t}) < \frac{\lfloor k \cdot n \rfloor}{n} \text{ and } \frac{\lfloor k \cdot n \rfloor}{n} < \frac{1}{n} \sum_i R_i\right\}$$

$$\geq Pr_{P^n}\left\{\widehat{B}_n(\overline{t}) < \frac{\lfloor k \cdot n \rfloor}{n}\right\} + Pr_{P^n}\left\{\frac{\lfloor k \cdot n \rfloor}{n} < \frac{1}{n} \sum_i R_i\right\} - 1 \qquad (24)$$

For the first term in (24), we have the following lower bound

$$Pr_{P^n} \left\{ \frac{1}{n} \sum_i R_i \cdot \mathbf{1}\{X_i \leq \bar{t}\} \leq \frac{k \cdot n - 1}{n} \right\}$$

$$= Pr_{P^n} \left\{ \sqrt{n} \left(\frac{1}{n} \sum_i R_i \cdot \mathbf{1}\{X_i \leq \bar{t}\} - k \right) \leq -\frac{1}{\sqrt{n}} \right\} \rightarrow 0.5$$

To see the convergence, we apply the Central Limit Theorem to the LHS, and note that $-\frac{1}{\sqrt{n}}$ converges to zero. Denote $p_R = Pr\{R = 1\}$. For the second term in (24), we have the following lower bound

$$Pr_{P^n} \left\{ \frac{1}{n} \sum_i R_i \geq \frac{k \cdot n}{n} \right\}$$

$$= Pr_{P^n} \left\{ \sqrt{n} \left(\frac{1}{n} \sum_i R_i - p_R \right) \geq \sqrt{n} \cdot (k - p_R) \right\} \rightarrow 1$$

To see the convergence, we apply the Central Limit Theorem to the LHS, and note that $\sqrt{n} \cdot (k - p_R)$ diverges to $-\infty$ for $p_R > k$. We thus conclude

$$\lim_{n \to \infty} Pr_{P^n} \left\{ B(\widehat{t}; P) > B(\bar{t}; P) \right\} \geq 0.5$$

which proves $\widehat{t}$ is not pointwise asymptotically feasible under the distribution P.

Similar argument shows $Pr_{P^n} \left\{ \widehat{B}_n(t^*) > \widehat{B}_n(\underline{t}) \right\}$ has a limit of one half. We thus conclude

$$\lim_{n \to \infty} Pr_{P^n} \left\{ W(\widehat{t}; P) < W(\underline{t}; P) \right\} \geq 0.5$$

which proves that $\widehat{t}$ is not pointwise asymptotically optimal under the distribution P. $\qquad \square$

Proof. **Proof of Theorem 3.1.** By construction, the limit probability for any policy in $\hat{\mathcal{G}}_\alpha$ to violate the budget constraint is

$$Pr_{P^n}\{\exists g : g \in \hat{\mathcal{G}}_\alpha \text{ and } B(g; P) > k\} = Pr_{P^n}\{\min_{B(g;P)>k} \frac{\sqrt{n}\left(\widehat{B}_n(g) - k\right)}{\widehat{\Sigma}^B(g,g)^{1/2}} \leq c_\alpha\}$$

$$\leq Pr_{P^n}\{\min_{B(g;P)>k} \frac{\sqrt{n}\left(\widehat{B}_n(g) - B(g; P)\right)}{\widehat{\Sigma}^B(g,g)^{1/2}} \leq c_\alpha\}$$

$$\leq Pr_{P^n}\{\min_{g\in\mathcal{G}} \frac{\sqrt{n}\left(\widehat{B}_n(g) - B(g; P)\right)}{\widehat{\Sigma}^B(g,g)^{1/2}} \leq c_\alpha\}$$

Under Assumption 2.4, uniformly over $P \in \mathcal{P}$, the empirical process $\left\{\sqrt{n}\left(\widehat{B}_n(g) - B(g; P)\right)\right\}$ converges to a Gaussian process G_P^B for $G_P^B(\cdot) \sim \mathcal{GP}(0, \Sigma_P^B(\cdot, \cdot))$ and we have a consistent covariance estimate $\widehat{\Sigma}^B(\cdot, \cdot)$. Then by the definition of c_α, we have

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} Pr_{P^n}\{\min_{g\in\mathcal{G}} \frac{\sqrt{n}\left(\widehat{B}_n(g) - B(g; P)\right)}{\widehat{\Sigma}^B(g,g)^{1/2}} \leq c_\alpha\}$$

$$= \sup_{P\in\mathcal{P}} Pr_{P^n}\{\inf_{g\in\mathcal{G}} \frac{G_P^B(g)}{\Sigma_P^B(g,g)^{1/2}} \leq c_\alpha\} = \alpha.$$

$\square$

Proof. **Proof of Theorem 3.2 and Corollary** 3.1. Theorem 3.2 follows from Theorem 3.1, replacing α with α_n in its proof. Specifically, by construction we have

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} Pr_{P^n}\{B(\hat{g}_{prob}; P) > k\} \leq \limsup_{n\to\infty} \sup_{P\in\mathcal{P}} Pr_{P^n}\{\exists g : g \in \hat{\mathcal{G}}_{\alpha_n} \text{ and } B(g; P) > k\} \leq \alpha_n.$$

For Corollary 3.1, we decompose the welfare loss into

$$W(g_P^*; P) - W(\hat{g}_{prob}; P)$$

$$= W(g_P^*; P) - \widehat{W}_n(\hat{g}_{prob}) + \widehat{W}_n(\hat{g}_{prob}) - W(\hat{g}_{prob}; P)$$

$$\leq \sup_{g\in\mathcal{G}} |W(g; P) - \widehat{W}_n(g)| + \widehat{W}_n(g_P^*) - \widehat{W}_n(\hat{g}_{prob})$$

Under the event that $g_P^* \in \hat{\mathcal{G}}_{\alpha_n}$, we are guaranteed that the last term is non-positive. This happens with probability

$$Pr_{P^n}\{\frac{\sqrt{n}\left(\widehat{B}_n(g_P^*) - k\right)}{\widehat{\Sigma}^B(g_P^*, g_P^*)^{1/2}} \leq c_{\alpha_n}\} \overset{a}{\sim} \Phi\left(c_{\alpha_n} + \frac{\sqrt{n}\left(k - B(g_P^*; P)\right)}{\Sigma_P^B(g_P^*, g_P^*)^{1/2}}\right) \to 1$$

for $B(g_P^*; P)$ strictly below the threshold. Since c_{α_n} diverges to $-\infty$ at a rate slower $\sqrt{n}$, the term in the parenthesis diverges to ∞ as $n \to \infty$. We then apply Lemma B.4 and B.6 (proved in Section B.3 under primitive conditions that lead to Assumption 2.4), which shows $\sup_{g \in \mathcal{G}} |W(g; P) - \widehat{W}_n(g)|$ converges to zero in probability. We then conclude $\hat{g}_{prob}$ is asymptotically optimal under distribution P. □

Proof. **Proof of Lemma 4.1.** By definition

$$W(g_P^*; P) = \max_{B(g;P) \leq k} W(g; P) = \max_{B(g;P) \leq k} W(g; P) - \overline{\lambda} \cdot (B(g; P) - k)_+$$

$$\leq \max_{g \in \mathcal{G}} W(g; P) - \overline{\lambda} \cdot (B(g; P) - k)_+$$

$$= W(\widetilde{g}_P; P) - \overline{\lambda} \cdot (B(\widetilde{g}_P; P) - k)_+ \leq W(\widetilde{g}_P; P)$$

$\square$

Proof. **Proof of Theorem 4.1.** We first prove $\widehat{g}_r$ is asymptotically optimal. Denote the new objective function with $V(g; P) = W(g; P) - \overline{\lambda} \cdot (B(g; P) - k)_+$, whose maximizer is $\widetilde{g}_P$. Denote $\widehat{V}_n(g) = -\widehat{W}_n(g) + \overline{\lambda} \cdot (\widehat{B}_n(g) - k)_+$ to be the sample-analog, whose maximizer is $\widehat{g}_r$. Apply uniform deviation bound to the difference between the value under $\widehat{g}_r$ and $\widetilde{g}_P$, we have

$$V(\widetilde{g}_P; P) - V(\widehat{g}_r; P)$$

$$= V(\widetilde{g}_P; P) - \widehat{V}_n(\widehat{g}_r) + \widehat{V}_n(\widehat{g}_r) - V(\widehat{g}_r; P)$$

$$\leq V(\widetilde{g}_P; P) - \widehat{V}_n(\widetilde{g}_P) + \widehat{V}_n(\widehat{g}_r) - V(\widehat{g}_r; P)$$

$$\leq 2 \sup_g \left| V(g; P) - \widehat{V}_n(g) \right|$$

$$= 2 \sup_g \left| W(g; P) - \widehat{W}_n(g) - \overline{\lambda} \cdot (\max\{B(g; P) - k, 0\} - \max\{\widehat{B}_n(g) - k, 0\}) \right|$$

$$\leq 2 \sup_g \left| \widehat{W}_n(g) - W(g; P) \right| + 2\overline{\lambda} \cdot \sup_g \left| \max\{\widehat{B}_n(g) - k, 0\} - \max\{B(g; P) - k, 0\} \right|$$

$$\leq 2 \sup_g \left| \widehat{W}_n(g) - W(g; P) \right| + 2\overline{\lambda} \cdot \sup_g \left| \widehat{B}_n(g) - B(g; P) \right|$$

The last line uses the fact that $|\max\{a, 0\} - \max\{b, 0\}| \leq |a - b|$. Both terms, $\sup_g \left| \widehat{W}_n(g) - W(g; P) \right|$ and $\sup_g \left| \widehat{B}_n(g) - B(g; P) \right|$ converge to zero in probability under Assumption 2.4. Specifically, Lemma B.4 and B.6 (proved in Section B.3 under primitive conditions that lead to Assumption 2.4) imply the uniform convergence

in probability

$$\sup_{P \in \mathcal{P}} \sup_{g \in \mathcal{G}} \left| \widehat{W}_n(g) - W(g; P) \right| \to_p 0, \ \sup_{P \in \mathcal{P}} \sup_{g \in \mathcal{G}} \left| \widehat{B}_n(g) - B(g; P) \right|_p \to 0.$$

We thus conclude

$$\sup_{P \in \mathcal{P}} \left| V(\widehat{g}_r; P) - V(\widetilde{g}_P; P) \right| \to_p 0.$$

Furthermore, we have $W(g_P^*; P) \leq V(\widetilde{g}_P; P) \leq W(\widetilde{g}_P; P)$ as a by-product of Lemma 4.1 for any $P \in \mathcal{P}$. Putting these together, we have

$$\inf_{P \in \mathcal{P}} \left\{ W(\widehat{g}_r; P) - W(g_P^*; P) \right\}$$

$$\geq \inf_{P \in \mathcal{P}} \left\{ V(\widehat{g}_r; P) - W(g_P^*; P) \right\}$$

$$\geq \inf_{P \in \mathcal{P}} \left\{ V(\widehat{g}_r; P) - V(\widetilde{g}_P; P) \right\} + \inf_{P \in \mathcal{P}} \left\{ V(\widetilde{g}_P; P) - W(g_P^*; P) \right\}$$

$$\geq \inf_{P \in \mathcal{P}} \left\{ V(\widehat{g}_r; P) - V(\widetilde{g}_P; P) \right\} \to_p 0.$$

which proves uniform asymptotic optimality. $\square$

B   Primitive assumptions and auxiliary lemmas

I first prove the optimal rule that solves the population constrained optimization problem takes the form of threshold. In Section B.2, I first provide primitive assumptions on the class of DGPs. I then prove in Lemma B.1, which establishes that these primitive assumptions imply Assumption 2.1.

In Section B.3, I verify Assumption 2.4 for settings where the observed sample comes from an RCT or an observational study, and the propensity score can be estimated efficiently based on parametric regressions.

B.1   Optimal rule without functional form restriction

The population problem is to find rules based on X_i that solves

$$\max_{g:\mathcal{X}\to\{0,1\}} E[\Gamma_i g(X_i)] \ \ s.t. \ E[R_i g(X_i)] < k$$

By Law of Iterated Expectation, we can write the constrained optimization problem as

$$\max_{g:\mathcal{X}\to\{0,1\}} E[\gamma(X_i)g(X_i)] \ \ s.t. \ E[r(X_i)g(X_i)] < k$$

where $\gamma(X_i) = E[\Gamma_i \mid X_i]$ and $r(X_i) = E[R_i \mid X_i]$.

Claim B.1. Let $d\mu = r(x)f(x)dx$ denote the positive measure. The constrained optimization problem is equivalent to

$$\max_{g:\mathcal{X}\to\{0,1\}} \int \frac{\gamma(x)}{r(x)} g(x) d\mu \ \ s.t. \ \int g(x) d\mu = k$$

Let X^* be the support of the solution g^*. It will take the form of $X^* = \{x : \frac{\gamma(x)}{r(x)} > c\}$ where c is chosen so that $\mu(X^*) = k$.

Proof. Let X be the support of any $g \neq g^*$ with $\mu(X) = k$. Then the objective

function associated g is

$$\int_{X^*} \frac{\gamma}{r} d\mu - \int_X \frac{\gamma}{r} d\mu = \int \frac{\gamma(x)}{r(x)} \mathbf{1}\{x \in X^*\} d\mu - \int \frac{\gamma(x)}{r(x)} \mathbf{1}\{x \in X\} d\mu$$

$$= \int \frac{\gamma(x)}{r(x)} \left(\mathbf{1}\{x \in X^* \backslash X\} - \mathbf{1}\{x \in X \backslash X^*\} \right) d\mu$$

By definition of X^*, we have $\frac{\gamma(x)}{r(x)} > c$ for $x \in X^* \backslash X$ and $\frac{\gamma(x)}{r(x)} < c$ for $x \in X \backslash X^*$. Also note that μ is a positive measure. Then the above difference is lower bounded by

$$\int \frac{\gamma(x)}{r(x)} \left(\mathbf{1}\{x \in X^* \backslash X\} - \mathbf{1}\{x \in X \backslash X^*\} \right) d\mu \geq c \int \left(\mathbf{1}\{x \in X^* \backslash X\} - \mathbf{1}\{x \in X \backslash X^*\} \right) d\mu \geq 0$$

as by construction, we have $\int \mathbf{1}\{x \in X^* \backslash X\} d\mu = 1$ since $\mu(X^*) = \mu(X) = k$. $\qquad \square$

B.2   Primitive assumptions for contiguity

Assumption B.1. *Assume the class of DGPs $\{P_\theta : \theta \in \Theta\}$ has densities p_θ with respect to some measure μ. Assume P_θ is DQM at P_0 i.e. $\exists \dot{\ell}_0$ s.t. $\int [\sqrt{p_h} - \sqrt{p_0} - \frac{1}{2} h' \dot{\ell}_0 \sqrt{p_0}]^2 d\mu = o(\|h^2\|)$ for $h \to 0$.*

Assumption B.2. *For all policies g, $B(g; P_\theta)$ is twice continuously differentiable in θ at 0, and the derivatives are bounded from above and away from zero within an open neighborhood $\mathcal{N}_\theta$ of zero uniformly over $g \in \mathcal{G}$.*

Lemma B.1. *Under Assumption B.1, the class $\mathcal{P}$ includes a sequence of data distribution $\{P_{h_n}\}$ that is contiguous to P_0 for every h_n satisfying $\sqrt{n} h_n \to h$ e.g. take $h_n = h/\sqrt{n}$. This proves the first part of Assumption 2.1. Suppose further Assumption B.2 holds, then there exists some h for the second part of Assumption 2.1 to hold,.*

Proof. **Proof of Lemma B.1.** By Theorem 7.2 of Vaart (1998), the log likelihood

ratio process converges under P_0 (denoted with $\overset{p_0}{\rightsquigarrow}$) to a normal experiment

$$\log \prod_{i=1}^{n} \frac{p_{h_n}}{p_0}(A_i) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h' \dot{\ell}_0(A_i) - \frac{1}{2} h' I_0 h + o_{P_0}(1) \tag{25}$$

$$\overset{p_0}{\rightsquigarrow} \mathcal{N}\left(-\frac{1}{2} h' I_0 h, h' I_0 h\right) \tag{26}$$

where $\dot{\ell}_0$ is the score and $I_{P_0} = E_{P_0}[\dot{\ell}_0(A_i)\dot{\ell}_0(A_i)']$ exists. The convergence in distribution of the log likelihood ratio to a normal with mean equal to $-\frac{1}{2}$ of its variance in (26) implies mutual contiguity $P_0^n \lhd\rhd P_{h_n}^n$ by Le Cam's first lemma (see Example 6.5 of Vaart (1998)). This proves the first part of the lemma.

By Taylor's theorem with remainder we have for each policy g

$$B(g; P_{h/\sqrt{n}}) - B(g; P_0) = \frac{h'}{\sqrt{n}} \frac{\partial B(g; P_0)}{\partial \theta} + \frac{1}{2} \frac{1}{n} h' \frac{\partial^2 B(g; P_{\tilde{\theta}_n})}{\partial \theta \partial \theta'} h$$

where $\tilde{\theta}_n$ is a sequence of values with $\tilde{\theta}_n \in [0, h/\sqrt{n}]$ that can depend on g. Take h so that the first term is positive for policies with $B(g; P_0) = k$. Such h exists because we assume $\frac{\partial B(g; P_0)}{\partial \theta}$ is bounded away from zero. For $g \in \mathcal{G}_0$ where $\mathcal{G}_0 = \{g : B(g; P_0) = k\}$, the constraints are violated under P_{h_n} and furthermore (multiplying by $\sqrt{n}$)

$$\sqrt{n} \cdot (B(g; P_{h_n}) - k) > h' \frac{\partial B(g; P_0)}{\partial \theta} > 0$$

for every n. This proves the second part of the lemma for $C = \inf_{g \in \mathcal{G}_0} \left| h' \frac{\partial B(g; P_0)}{\partial \theta} \right|$

. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

B.3 Primitive assumptions and proofs for estimation quality

In this section, I verify that $\widehat{W}_n(g)$ and $\widehat{B}_n(g)$ satisfy Assumption 2.4 under primitive assumptions on the policy class and the OHIE.

Assumption B.3. *VC-class: The policy class $\mathcal{G}$ has a finite VC-dimension $v < \infty$.*

To introduce the assumptions on the OHIE, I first recall the definitions of components in $\widehat{W}_n(g)$ and $\widehat{B}_n(g)$:

$$\widehat{W}_n(g) := \frac{1}{n}\sum_i \Gamma_i^* \cdot g(X_i), \ \ \widehat{B}_n(g) := \frac{1}{n}\sum_i R_i^* \cdot g(X_i)$$

for the doubly-robust scores

$$\Gamma_i^* = \widehat{\gamma}^Y(V_i, 1) - \widehat{\gamma}^Y(V_i, 0) + \widehat{\alpha}(V_i, D_i) \cdot \left(Y_i - \widehat{\gamma}^Y(V_i, D_i)\right)$$
$$R_i^* = \widehat{\gamma}^Z(V_i, 1) + \frac{D_i}{\widehat{p}(V_i)} \cdot \left(Z_i - \widehat{\gamma}^Z(V_i, D_i)\right)$$

and observed characteristics X_i. Here V_i collects the confounders in OHIE, namely household size (number of adults entered on the lottery sign-up form) and survey wave. Note that while X_i can overlap with V_i, the policy $g(X_i)$ needs not vary by V_i. In the OHIE example, the policy is based on number of children and income. However, conditional on household size and survey wave, income and number of children is independent of the lottery outcome in OHIE.

Recall that $\widehat{W}_n(g)$ and $\widehat{B}_n(g)$ are supposed to approximate net benefit Γ and net excess cost R of Medicaid eligibility. I provide more precise definitions for Γ and R as the primitive assumptions are stated in terms of their components.

Let $Y(1)$ be the (potential) subjective health when one is given Medicaid eligibility, and $Y(0)$ be the (potential) subjective health when one is not given Medicaid eligibility. Recall the definition for

$$\Gamma = Y(1) - Y(0)$$

We only observe the actual subjective health Y_i.

Let $M(1)$ be the (potential) enrollment in Medicaid when one is given Medicaid eligibility, and $M(0)$ be the (potential) enrollment in Medicaid when one is not given Medicaid eligibility.

Let $C(1)$ be the (potential) cost to the government when one is given Medicaid eligibility, and $C(0)$ be the (potential) cost to the government when one is not given Medicaid eligibility. Note that $C(0) = 0$ by construction. Even when given eligibility, one might not enroll and thus incur zero cost to the government. So given the current expenditure is \$6,000 per *enrollee*, the implied expenditure *per capita under eligibility policy $g(X)$* is actually $\$6,000 \cdot E[M(1)g(X)]$. The reason is that the expected enrollment rate is only $Pr\{M(1) = 1 \mid X \in g\}$. So the per capita *excess* cost of Medicaid eligibility policy $g(X)$ relative to the current level is

$$R = C(1) - \$6,000 \cdot M(1).$$

We only observe the actual cost $(C_i = D_i C_i(1) + (1 - D_i)C_i(0))$ and the actual Medicaid enrollment (M_i) in OHIE. We calculate $Z_i = C_i - \$6,000 \cdot M_i$.

Assumption B.4. *Suppose for all $P \in \mathcal{P}$, the following statements hold for the OHIE:*

Independent characteristics: $Pr\{D_i = 1 \mid V_i, X_i\} = Pr\{D_i = 1 \mid V_i\}$

Unconfoundedness: $(Y(1), Y(0), C(1), M(1)) \perp D_i | V_i$.

Bounded attributes: the support of variables X_i, Y_i and Z_i are bounded.

Strict overlap: There exist $\kappa \in (0, 1/2)$ such that the propensity score satisfies $p(v) \in [\kappa, 1 - \kappa]$ for all $v \in \mathcal{V}$.

B.3.1 Uniform convergence of $\widehat{W}_n(\cdot)$ and $\widehat{B}_n(\cdot)$

We want to show the recentered empirical processes $\widehat{W}_n(\cdot)$ and $\widehat{B}_n(\cdot)$ converge to mean-zero Gaussian processes G_P^W and G_P^B with covariance functions $\Sigma_P^W(\cdot, \cdot)$ and $\Sigma_P^B(\cdot, \cdot)$ respectively uniformly over $P \in \mathcal{P}$. The covariance functions are uniformly bounded, with diagonal entries bounded away from zero uniformly over $g \in \mathcal{G}$. Take $\widehat{W}_n(\cdot)$ for example, the recentered empirical processes is

$$\sqrt{n}\left(\frac{1}{n}\sum_i \Gamma_i^* \cdot g(X_i) - E_P[\Gamma \cdot g(X_i)]\right)$$

and can be expressed as the sum of two terms

$$\frac{1}{\sqrt{n}}\sum_i(\Gamma_i^* - \widetilde{\Gamma}_i) \cdot g(X_i) + \sqrt{n}\left(\frac{1}{n}\sum_i \widetilde{\Gamma}_i \cdot g(X_i) - E_P[\Gamma \cdot g(X_i)]\right). \qquad (27)$$

Here $\widetilde{\Gamma}_i$ are the theoretical analogs

$$\widetilde{\Gamma}_i = \gamma^Y(V_i, 1) - \gamma^Y(V_i, 0) + \alpha(V_i, D_i) \cdot \left(Y_i - \gamma^Y(V_i, D_i)\right)$$

which is doubly-robust score with the theoretical propensity score and the CEF. A similar expansion holds for $\widehat{B}_n(\cdot)$ involving the theoretical analog

$$\widetilde{R}_i = \gamma^Z(V_i, 1) + \frac{D_i}{p(V_i)} \cdot \left(Z_i - \gamma^Z(V_i, D_i)\right).$$

The following lemmas prove the uniform convergence of $\widehat{W}_n(\cdot)$ and $\widehat{B}_n(\cdot)$.

The last part of the assumption is that we have a uniformly consistent estimate for the covariance function. I argue the sample analog

$$\widehat{\Sigma}^B(g, g') = \frac{1}{n}\sum_i(\widetilde{R}_i)^2 \cdot g(X_i) \cdot g'(X_i) - \left(\frac{1}{n}\sum_i \widetilde{R}_i \cdot g(X_i)\right) \cdot \left(\frac{1}{n}\sum_i \widetilde{R}_i \cdot g'(X_i)\right)$$

is a pointwise consistent estimate for $\widehat{\Sigma}^B(g, g')$. To see this, note that the covariance function $\Sigma_P^B(\cdot, \cdot)$ maps $\mathcal{G} \times \mathcal{G}$ to $\mathbb{R}$. Under Assumption B.4 that $\mathcal{G}$ is a VC-class, the product $\mathcal{G} \times \mathcal{G}$ is also a VC-class (see e.g. Lemma 2.6.17 of van der Vaart and Wellner (1996)). By a similar argument leading to Lemma B.4 and B.6 , we conclude the uniform consistency of $\widehat{\Sigma}^B(\cdot, \cdot)$.

Lemma B.2. *Let $\mathcal{G}$ be a VC-class of subsets of $\mathcal{X}$ with VC-dimension $v < \infty$. The*

following sets of functions from $\mathcal{A}$ to $\mathbb{R}$

$$\mathcal{F}^W = \{\widetilde{\Gamma}_i \cdot g(X_i) : g \in \mathcal{G}\}$$

$$\mathcal{F}^B = \{\widetilde{R}_i \cdot g(X_i) : g \in \mathcal{G}\}$$

are VC-subgraph class of functions with VC-dimension less than or equal to v for all $P \in \mathcal{P}$. For notational simplicity, we suppress the dependence of $\mathcal{F}$ on P.

Lemma B.3. *For all P in the family $\mathcal{P}$ of distributions satisfying Assumption B.4, for all $g \in \mathcal{G}$ we have*

$$E_P[\widetilde{\Gamma}_i \cdot g(X_i)] = W(g; P)$$

$$E_P[\widetilde{R}_i \cdot g(X_i)] = B(g; P)$$

Lemma B.4. *Let $\mathcal{G}$ satisfy Assumption B.4 (VC). Let U_i be a mean-zero bounded random vector of fixed dimension i.e. there exists $M < \infty$ such that i.e. $U_i \in [-M/2, M/2]$ almost surely under all $P \in \mathcal{P}$. Then the uniform deviation of the sample average of vanishes*

$$\sup_{P \in \mathcal{P}} \sup_{g \in \mathcal{G}} \left| \frac{1}{n} \sum_i U_i \cdot g(X_i) \right| \to_{a.s.} \mathbf{0}.$$

Lemma B.5. *Let $\mathcal{P}$ be a family of distributions satisfying Assumption B.4. Let $\mathcal{G}$ satisfy Assumption B.4 (VC). Then $\mathcal{F}^W$ and $\mathcal{F}^B$ are P-Donsker for all $P \in \mathcal{P}$. That is, the empirical process indexed by $g \in \mathcal{G}$*

$$\sqrt{n} \cdot \left(\frac{1}{n} \sum_i \widetilde{\Gamma}_i \cdot g(X_i) - W(g; P) \right)$$

converge to a Gaussian process $\mathcal{GP}(0, \Sigma_P^W(\cdot, \cdot))$ uniformly in $P \in \mathcal{P}$, and the empirical

process indexed by $g \in \mathcal{G}$

$$\sqrt{n} \cdot (\frac{1}{n} \sum_i \widetilde{R}_i \cdot g(X_i) - B(g; P))$$

converge to a Gaussian process $\mathcal{GP}(0, \Sigma_P^B(\cdot, \cdot))$ uniformly in $P \in \mathcal{P}$.

Lemma B.6. *Let $\mathcal{P}$ be a family of distributions satisfying Assumption B.4. Let $\mathcal{G}$ satisfy Assumption B.4 (VC). Then the estimation errors vanishes*

$$\sup_{P \in \mathcal{P}} \sup_{g \in \mathcal{G}} \left| \frac{1}{\sqrt{n}} \sum_i (\Gamma_i^* - \widetilde{\Gamma}_i) \cdot g(X_i) \right| \to_p 0$$

$$\sup_{P \in \mathcal{P}} \sup_{g \in \mathcal{G}} \left| \frac{1}{\sqrt{n}} \sum_i (R_i^* - \widetilde{R}_i) \cdot g(X_i) \right| \to_p 0$$

B.3.2 Proofs of auxiliary lemmas

Proof. **Proof of Lemma B.2.** This lemma follows directly from Lemma A.1 of Kitagawa and Tetenov (2018). □

Proof. **Proof of Lemma B.3.** Under Assumption B.4, we prove each $\widetilde{\Gamma}_i$ is an conditionally unbiased estimate for Γ:

$$E_P[\widetilde{\Gamma}_i g(X_i)] = E_P[E_P[\widetilde{\Gamma}_i \mid X_i] g(X_i)] = E_P[E_P[\Gamma \mid X_i] g(X_i)] = W(g; P).$$

We focus on $E_P[\widetilde{\Gamma}_i \mid X_i] = E_P[E_P[\widetilde{\Gamma}_i \mid V_i, X_i] \mid X_i]$. However, conditional on V_i, by unconfoundedness and strict overlap we have

$$E_P[\widetilde{\Gamma}_i \mid V_i, X_i] = E[Y_i \mid V_i, X_i, D_i = 1] - E[Y_i \mid V_i, X_i, D_i = 0]$$

$$= E[Y_i(1) - Y_i(0) \mid V_i, X_i] = E[\Gamma_i \mid V_i, X_i]$$

Specifically

$$E[\widetilde{\Gamma}_i \mid V_i, X_i] = E\left[\gamma^Y(V_i, 1) - \gamma^Y(V_i, 0) \mid V_i, X_i\right] + E\left[\alpha(V_i, D_i) \cdot \left(Y_i - \gamma^Y(V_i, D_i)\right) \mid V_i, X_i\right]$$

$$= E\left[Y_i \mid V_i, D_i = 1\right] - E\left[Y_i \mid V_i, D_i = 0\right] +$$

$$E\left[Y_i(1) - Y_i(0) \mid V_i, X_i\right] - \left(E\left[Y_i \mid V_i, D_i = 1\right] - E\left[Y_i \mid V_i, D_i = 0\right]\right)$$

$$= E[\Gamma \mid V_i, X_i]$$

Specifically, we expand $E[\alpha(V_i, D_i)Y_i \mid V_i, X_i]$

$$= E\left[\frac{Y_i}{p(V_i)} \mid D_i = 1, V_i, X_i\right] \cdot Pr\{D_i = 1 \mid V_i, X_i\} - E\left[\frac{Y_i}{1 - p(V_i)} \mid D_i = 0, X_i\right] \cdot Pr\{D_i = 0 \mid V_i, X_i\}$$

$$= E\left[D_iY_i(1) + (1 - D_i)Y_i(0) \mid D_i = 1, V_i, X_i\right] - E\left[D_iY_i(1) + (1 - D_i)Y_i(0) \mid D_i = 0, V_i, X_i\right]$$

$$= E\left[Y_i(1) \mid V_i, X_i\right] - E\left[Y_i(0) \mid V_i, X_i\right]$$

where the second line holds by independent characteristic such that

$$Pr\{D_i = 1 \mid V_i, X_i\} = Pr\{D_i = 1 \mid V_i\} =: p(V_i)$$

and the last line follows from unconfoundedness. Similarly, we show

$$E\left[\alpha(V_i, D_i) \cdot \gamma^Y(V_i, D_i) \mid V_i, X_i\right] = E\left[Y_i \mid V_i, D_i = 1\right] - E\left[Y_i \mid V_i, D_i = 0\right]$$

Similar argument holds for $E_P[\widetilde{R}_i \cdot g(X_i)]$, with the only modification:

$$E[\frac{D_i}{p(V_i)}C_i \mid V_i, X_i] = E\left[\frac{C_i}{p(V_i)} \mid D_i = 1, V_i, X_i\right] \cdot Pr\{D_i = 1 \mid V_i, X_i\}$$

$$= E[C_i(1) \mid V_i, X_i]$$

$\square$

Proof. **Proof of Lemma B.4.** Denote the following set of functions from $\mathcal{U}$ to $\mathbb{R}$

$$\mathcal{F}^U = \{U_i \cdot g(X_i) : g \in \mathcal{G}\}$$

and it has uniform envelope $\bar{F} = M/2$ since U_i is bounded. This envelop function is bounded uniformly over $\mathcal{P}$. Also, by Assumption B.4 (VC) and Lemma B.2, $\mathcal{F}^U$ is VC-subgraph class of functions with VC-dimension at most v. By Lemma 4.14 and Proposition 4.18 of Wainwright (2019), we conclude that $\mathcal{F}^U$ has Rademacher complexity $2\sqrt{M^2 \frac{v}{n}}$. Then by Proposition 4.12 of Wainwright (2019) we conclude that $\mathcal{F}^U$ are P-Glivenko–Cantelli for each $P \in \mathcal{P}$, with an $O(\sqrt{\frac{v}{n}})$ rate of convergence. Note that this argument does not use any constants that depend on P but only M and v, so we can actually get uniform convergence over $\mathcal{P}$. $\qquad\square$

Proof. **Proof of Lemma B.5.** Note that Assumption B.4 imply that $\mathcal{F}^W$ and $\mathcal{F}^B$ have uniform envelope $\bar{F} = M/(2\kappa)$. $\mathcal{F}^W$ and $\mathcal{F}^B$ thus have square integrable envelop functions uniformly over $\mathcal{P}$. Also, by Assumption B.4 (VC) and Lemma B.2, $\mathcal{F}^W$ and $\mathcal{F}^B$ are VC-subgraph class of functions with VC-dimension at most v. Even though both $\mathcal{F}^W$ and $\mathcal{F}^B$ depend on P, a similar argument for Theorem 1 in Rai (2019) show that $\mathcal{F}^W$ and $\mathcal{F}^B$ are P-Donsker uniformly in $P \in \mathcal{P}$. $\qquad\square$

Proof. **Proof of Lemma B.6.** We focus on the deviation in $\Gamma_i^* - \widetilde{\Gamma}_i$. The deviation in $R_i^* - \widetilde{R}_i$ can be proven to vanish in a similar manner. Denote $\Delta \gamma^Y(V_i) = \gamma^Y(V_i, 1) - \gamma^Y(V_i, 0)$. For any fixed policy g, we expand the deviation $\frac{1}{\sqrt{n}} \sum_i (\Gamma_i^* - \widetilde{\Gamma}_i) g(X_i)$ into three terms

$$\frac{1}{\sqrt{n}} \sum_i g(X_i) \left(\widehat{\alpha}(V_i, D_i) - \alpha(V_i, D_i)\right) \cdot \left(Y_i - \gamma^Y(V_i, D_i)\right) \tag{28}$$

$$+ \frac{1}{\sqrt{n}} \sum_i g(X_i) \left(\Delta \widehat{\gamma}^Y(V_i) - \Delta \gamma^Y(V_i) - \alpha(V_i, D_i) \cdot \left(\widehat{\gamma}^Y(V_i, D_i) - \gamma^Y(V_i, D_i)\right)\right)$$

$$- \frac{1}{\sqrt{n}} \sum_i g(X_i) \left(\widehat{\gamma}^Y(V_i, D_i) - \gamma^Y(V_i, D_i)\right) \cdot \left(\widehat{\alpha}(V_i, D_i) - \alpha(V_i, D_i)\right) \cdot$$

Denote these three summands by $D_1(g)$, $D_2(g)$ and $D_3(g)$. We will bound all three summands separately. Recall we use the full sample to estimate the propensity score and the CEF with a saturated model. The purpose of the above expansion is to separately bound the estimation error from the estimated CEF and propensity score, and the deviation from taking sample averages. For cross-fitted estimators for the propensity score and the CEF, a similar bound can be found in Athey and Wager (2020).

Uniform consistency of the estimated CEF and propensity score Denote with $b(V_i, D_i)$ the dictionary that spans (V_i, D_i), and $b(V_i,)$ the dictionary that spans V_i. The saturated models are therefore parameterized as $\gamma^Z(V_i, D_i) = \gamma' b(V_i, D_i)$ and $p(V_i) = \beta' b(V_i)$. Under standard argument, the OLS estimators $\widehat{\gamma}$ and $\widehat{\beta}$ are asymptotically normal uniformly over $P \in \mathcal{P}$:

$$\sqrt{n} \cdot (\widehat{\gamma} - \gamma) = O_P(1), \quad \sqrt{n} \cdot \left(\widehat{\beta} - \beta\right) = O_P(1).$$

Furthermore, the in-sample L_2 errors from the estimated CEF and propensity score vanish. Consider

$$\frac{1}{n}\sum_i \left(\widehat{\gamma}^Z(V_i, D_i) - \gamma^Z(V_i, D_i)\right)^2 = \frac{1}{n}\sum_i \left((\widehat{\gamma} - \gamma)' b(V_i, D_i)\right)^2 = (\widehat{\gamma} - \gamma)' \widehat{M} (\widehat{\gamma} - \gamma)$$

where $\widehat{M} = \frac{1}{n}\sum_i b(V_i, D_i)b(V_i, D_i)'$. It converges in probability to a fixed matrix $M = E[b(V_i, D_i)b(V_i, D_i)']$. So the in-sample L_2 error from the estimated CEF vanishes at the rate of $n^{-1/2}$.

Similarly, consider expanding $\frac{1}{n}\sum_i \left(\widehat{\alpha}(V_i, D_i) - \alpha(V_i, D_i)\right)^2$ as

$$\frac{1}{n}\sum_i \left(\frac{1}{\widehat{\beta}'b(V_i)} - \frac{1}{\beta'b(V_i)}\right)^2 D_i^2 + \left(\frac{1}{1 - \widehat{\beta}'b(V_i)} - \frac{1}{1 - \beta'b(V_i)}\right)^2 (1 - D_i)^2$$

With a first-order Taylor approximation, for each term in the summand, the domi-

nating term would be

$$\frac{1}{n}\sum_i \left(\left(\widehat{\beta}-\beta\right)' \frac{-b(V_i)}{(\beta'b(V_i))^2} \right)^2 D_i^2 = \left(\widehat{\beta}-\beta\right)' \left(\frac{1}{n}\sum_i \frac{b(V_i)b(V_i)'}{(\beta'b(V_i))^2} D_i^2 \right) \left(\widehat{\beta}-\beta\right)$$

where the middle term converges to a fixed matrix as implied by $\beta'b(V_i)$ being bounded away from zero and one. So the in-sample L_2 error from the estimated propensity score also vanishes at the rate of $n^{-1/2}$.

Bounding the deviation We now bound each term in (28). Plugging in the first-order Taylor approximation with a remainder term to the estimated propensity score, we have

$$
\begin{aligned}
D_1(g) =& \frac{1}{\sqrt{n}}\sum_i g(X_i)\left(Y_i - \gamma^Y(V_i,D_i)\right) \cdot \\
& \left(\left(\frac{1}{\widehat{\beta}'b(V_i)} - \frac{1}{\beta'b(V_i)} \right) D_i + \left(\frac{1}{1-\widehat{\beta}'b(V_i)} - \frac{1}{1-\beta'b(V_i)} \right)(1-D_i) \right) \\
=& \frac{1}{\sqrt{n}}\sum_i g(X_i)\left(\widehat{\beta}-\beta\right)' \frac{-b(V_i)}{(\beta'b(V_i))^2} \cdot D_i \cdot \left(Y_i - \gamma^Y(V_i,D_i)\right) + \\
& \frac{1}{\sqrt{n}}\sum_i g(X_i)\left(\widehat{\beta}-\beta\right)' \frac{b(V_i)b(V_i)'}{(\widetilde{\beta}'V_i)^3}\left(\widehat{\beta}-\beta\right) \cdot D_i \cdot \left(Y_i - \gamma^Y(V_i,D_i)\right) \\
=& \underbrace{\sqrt{n}\left(\widehat{\beta}-\beta\right)'}_{O_P(1)} \underbrace{\frac{1}{n}\sum_i g(X_i)\frac{-b(V_i)}{(\beta'b(V_i))^2} \cdot D_i \cdot \left(Y_i - \gamma^Y(V_i,D_i)\right)}_{o_P(1)} + o_P(1)
\end{aligned}
$$

where $\widetilde{\beta}$ is a sequence between $\widehat{\beta}$ and β. This remainder term therefore converges to zero. Uniform convergence of the sample average follows from Lemma B.4: the random vector $\frac{-b(V_i)}{(\beta'b(V_i))^2} \cdot D_i \cdot \left(Y_i - \gamma^Y(V_i,D_i)\right)$ is mean-zero, has fixed dimension, and bounded.

We can decompose $D_2(g)$ into the product of two terms

$$D_2(g) = \underbrace{\sqrt{n}\,(\widehat{\gamma} - \gamma)}_{O_p(1)}{}' \frac{1}{n} \sum_i g(X_i)\left(\Delta b(V_i, D_i) - \alpha(V_i, D_i) \cdot b(V_i, D_i)\right)$$

Uniform convergence of the sample average again follows from Lemma B.4. We thus conclude $D_1(g)$ and $D_2(g)$ vanish uniformly over $g \in \mathcal{G}$ and over $P \in \mathcal{P}$.

For $D_3(g)$, we apply the Cauchy-Schwarz inequality to note that

$$D_3(g) \leq \sqrt{n} \cdot \sqrt{\frac{1}{n}\sum_i \left(\widehat{\gamma}^Y(V_i, D_i) - \gamma^Y(V_i, D_i)\right)^2} \cdot \sqrt{\frac{1}{n}\sum_i \left(\widehat{\alpha}(V_i, D_i) - \alpha(V_i, D_i)\right)^2}$$

The terms in the square root are the in-sample L_2 errors from the estimated CEF and propensity score, which vanish at the rate of $n^{-1/2}$ uniformly over $P \in \mathcal{P}$ as shown in the paragraph above. We thus conclude $D_3(g)$ vanishes uniformly over $g \in \mathcal{G}$ and over $P \in \mathcal{P}$. $\qquad\square$